```
OM protein - nucleic search, using frame_plus_p2n model

Run on:          March 30, 2003, 01:24:25 ; Search time 10977.4 Seconds
                                            (without alignments)
                                            2152.747 Million cell updates/sec

Title:           US-09-873-409-2
Perfect score:   4079
Sequence:        1 MVDENDIRALNVRHYRDHIG..........QELLRNRDIYFKLVNAQSVQ 812

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters:      4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115418_19228/app_query.fasta_1.7544
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_54887_@runat_27032003_115418_19228 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       GenEmbl:*
                 1:  gb_ba:*
                 2:  gb_htg:*
                 3:  gb_in:*
                 4:  gb_om:*
                 5:  gb_ov:*
                 6:  gb_pat:*
                 7:  gb_ph:*
                 8:  gb_pl:*
                 9:  gb_pr:*
                 10:  gb_ro:*
                 11:  gb_sts:*
                 12:  gb_sy:*
                 13:  gb_un:*
                 14:  gb_vi:*
                 15:  em_ba:*
                 16:  em_fun:*
                 17:  em_hum:*
                 18:  em_in:*
                 19:  em_mu:*
                 20:  em_om:*
                 21:  em_or:*
                 22:  em_ov:*
                 23:  em_pat:*
                 24:  em_ph:*
                 25:  em_pl:*
                 26:  em_ro:*
                 27:  em_sts:*
                 28:  em_un:*
                 29:  em_vi:*
                 30:  em_htg_hum:*
                 31:  em_htg_inv:*
                 32:  em_htg_other:*
                 33:  em_htg_mus:*
                 34:  em_htg_pln:*
                 35:  em_htg_rod:*
                 36:  em_htg_mam:*
                 37:  em_htg_vrt:*
                 38:  em_sy:*
                 39:  em_htgo_hum:*
                 40:  em_htgo_mus:*
                 41:  em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4079	100.0	2856	6	AX339028	AX339028 Sequence
2	4079	100.0	3177	6	AX339030	AX339030 Sequence
3	4079	100.0	3621	6	AX339032	AX339032 Sequence
4	4079	100.0	3702	6	AX339031	AX339031 Sequence
5	3919	96.1	3699	6	AX478104	AX478104 Sequence
6	3456	84.7	2066	6	AX339027	AX339027 Sequence
7	2480	60.8	4018	5	GGA9799	AJ009799 Gallus ga
8	2476	60.7	3987	10	CRUPGP1185	M59254 Chinese ham
9	2476	60.7	4296	10	CRUPGPI	M60040 C.griseus P
10	2476	60.7	4304	10	CRUPGP1165	M59253 Chinese ham
11	2468	60.5	3858	4	OAU78609	U78609 Ovis aries
12	2467.5	60.5	4084	10	MUSMDRA	J03398 Mouse mdr g
13	2465.5	60.4	3924	6	AX024455	AX024455 Sequence
14	2465.5	60.4	3924	6	AX409652	AX409652 Sequence
15	2465.5	60.4	3924	9	HUMMDR3	M23234 Human membr
16	2451	60.1	4927	10	AF257746	AF257746 Rattus no
17	2442	59.9	4279	6	AX105057	AX105057 Sequence
18	2442	59.9	4279	6	AX105078	AX105078 Sequence
19	2442	59.9	4279	6	AX105080	AX105080 Sequence
20	2441	59.8	3934	4	CFA419568	AJ419568 Canis fam
21	2441	59.8	4045	12	AF269224	AF269224 Synthetic
22	2441	59.8	4186	6	AX108654	AX108654 Sequence
23	2441	59.8	4195	6	AX108656	AX108656 Sequence
24	2441	59.8	4279	6	AX105082	AX105082 Sequence
25	2438	59.8	4317	4	AF045016	AF045016 Canis fam
26	2438	59.8	4317	6	AX105059	AX105059 Sequence
27	2438	59.8	4323	10	AF286167	AF286167 Rattus no
28	2437	59.7	2752	5	AF099732	AF099732 Fundulus
29	2436	59.7	3840	6	AX481416	AX481416 Sequence
30	2436	59.7	3860	6	AX322787	AX322787 Sequence
31	2436	59.7	3860	6	AX322789	AX322789 Sequence
32	2436	59.7	3988	6	AX024454	AX024454 Sequence
33	2436	59.7	4192	9	AF016535	AF016535 Homo sapi
34	2436	59.7	4264	6	AR051647	AR051647 Sequence
35	2436	59.7	4264	6	AR051650	AR051650 Sequence
36	2436	59.7	4646	6	AX336420	AX336420 Sequence
37	2436	59.7	4646	6	AX336708	AX336708 Sequence
38	2436	59.7	4646	6	AX391099	AX391099 Sequence
39	2436	59.7	4646	6	I49610	I49610 Sequence 2
40	2436	59.7	4646	9	HUMMDR1	M14758 Homo sapien
41	2436	59.7	4669	6	I08557	I08557 Sequence 3
42	2436	59.7	6505	6	AR028671	AR028671 Sequence
43	2436	59.7	8630	6	AX012320	AX012320 Sequence
44	2436	59.7	8630	6	AX012321	AX012321 Sequence
45	2436	59.7	9318	6	AR028672	AR028672 Sequence

ALIGNMENTS

RESULT 1

```
AX339028
LOCUS       AX339028                2856 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 10 from Patent WO0194400.
ACCESSION   AX339028
VERSION     AX339028.1  GI:18129120
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 10 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..2856
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      888 a    540 c    652 g    776 t
ORIGIN

Alignment Scores:
Pred. No.:              1.87e-296     Length:        2856
Score:                  4079.00       Matches:       812
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     6             Gaps:          0

US-09-873-409-2 (1-812) x AX339028 (1-2856)

Qy        1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      418 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 477

Qy       21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      478 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 537

Qy       41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      538 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 597

Qy       61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      598 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 657

Qy       81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      658 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 717

Qy      101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      718 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 777

Qy      121 AlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThr 140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      778 GCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTAGCACACCGACTTTCTACT 837

Qy      141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      838 ATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATGCTGGCGGAGAAAGGAGCA 897

Qy      161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      898 CATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTTGTGATGTCACAGGATATT 957

Qy      181 LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer 200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      958 AAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCA 1017

Qy      201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSer 220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1018 CTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCC 1077

Qy      221 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1078 ACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAAC 1137

Qy      241 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1138 AAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCTGTTCTAAATGGAACTGTT 1197

Qy      261 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1198 CATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATGTTTGGAAATAATGATAAA 1257

Qy      281 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1258 ACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTCGTCATTTTGGGTGTTATT 1317

Qy      301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1318 TGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGAGCAGGGGAAATTTTAACG 1377

Qy      321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1378 ATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAGGATATTGCCTGGTTTGAT 1437

Qy      341 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1438 GAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCCATAGATATAGCACAAATT 1497

Qy      361 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1498 CAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAATGCAACTAACATGGGACTT 1557

Qy      381 SerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAla 400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1558 TCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTCCTGATTCTGAGTATTGCT 1617

Qy      401 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1618 CCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATGACTGGATTTGCCAACAAA 1677

Qy      421 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1678 GATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAGAATATACGT 1737

Qy      441 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1738 ACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATGTATGAAGAGATGCTTCAG 1797

Qy      461 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1798 ACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTATGCATTCAGC 1857

Qy      481 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1858 CATGCCTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTTGGAGCCTATTTAATTCAA 1917

Qy      501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1918 GCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACTGCAATTGCATATGGAGCT 1977

Qy      521 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1978 ATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCCAAAGCCAAATCGGGGGCT 2037

Qy      541 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2038 GCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGACAGCCGCAGTCAAGAAGGG 2097

Qy      561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 580
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2098 AAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTCTTCTATCCA 2157

Qy       581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2158 TGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACA 2217

Qy       601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2218 GTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCTGTTCAACTTCTGCAGAGA 2277

Qy       621 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2278 CTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTGGATGCAAAAGAATTGAAT 2337

Qy       641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2338 GTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAGCCTGTGCTCTTCAACTGC 2397

Qy       661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2398 AGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTGGTGCCATTAGATGAGATC 2457

Qy       681 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2458 AAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAAGGTCTCCCTGAGAAATAC 2517

Qy       701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2518 AACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCT 2577

Qy       721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2578 ATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCC 2637

Qy       741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2638 CTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGATAAAGCCAGGACGGGAAGG 2697

Qy       761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2698 ACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGGTT 2757

Qy       781 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2758 CTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAGCTCCTGAGAAATCGAGAC 2817

Qy       801 IleTyrPheLysLeuValAsnAlaGlnSerValGln 812
             ||||||||||||||||||||||||||||||||||||
Db      2818 ATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 2853
```

```
RESULT 2
AX339030
LOCUS       AX339030              3177 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 12 from Patent WO0194400.
ACCESSION   AX339030
VERSION     AX339030.1  GI:18129122
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 12 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..3177
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      970 a    601 c    735 g    870 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:               2.16e-296      Length:         3177
Score:                   4079.00        Matches:        812
Percent Similarity:      100.00%        Conservative:   0
Best Local Similarity:   100.00%        Mismatches:     0
Query Match:             100.00%        Indels:         0
DB:                      6              Gaps:           0

US-09-873-409-2 (1-812) x AX339030 (1-3177)
```

```
Qy         1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       739 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 798

Qy        21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       799 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 858

Qy        41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       859 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 918

Qy        61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       919 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 978

Qy        81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       979 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 1038

Qy       101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1039 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 1098

Qy       121 AlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThr 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1099 GCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTAGCACACCGACTTTCTACT 1158

Qy       141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1159 ATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATGCTGGCGGAGAAAGGAGCA 1218

Qy       161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1219 CATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTTGTGATGTCACAGGATATT 1278

Qy       181 LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer 200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1279 AAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCA 1338

Qy       201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSer 220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1339 CTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCC 1398

Qy       221 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1399 ACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAAC 1458

Qy       241 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1459 AAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCTGTTCTAAATGGAACTGTT 1518

Qy       261 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1519 CATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATGTTTGGAAATAATGATAAA 1578

Qy       281 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1579 ACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTCGTCATTTTGGGTGTTATT 1638

Qy       301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 320
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1639 TGCTTTGTCAGTTATTTCATGCAGGGATTATTTACGGCAGAGCAGGGGAAATTTTAACG 1698

Qy       321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1699 ATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAGGATATTGCCTGGTTTGAT 1758

Qy       341 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1759 GAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCCATAGATATAGCACAAATT 1818

Qy       361 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1819 CAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAATGCAACTAACATGGGACTT 1878

Qy       381 SerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAla 400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1879 TCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTCCTGATTCTGAGTATTGCT 1938

Qy       401 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1939 CCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATGACTGGATTTGCCAACAAA 1998

Qy       421 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1999 GATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAGAATATACGT 2058

Qy       441 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2059 ACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATGTATGAAGAGATGCTTCAG 2118

Qy       461 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2119 ACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTATGCATTCAGC 2178

Qy       481 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2179 CATGCCTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTTGGAGCCTATTTAATTCAA 2238

Qy       501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2239 GCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACTGCAATTGCATATGGAGCT 2298

Qy       521 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2299 ATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCCAAAGCCAAATCGGGGGCT 2358

Qy       541 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2359 GCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGACAGCCGCAGTCAAGAAGGG 2418

Qy       561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2419 AAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTCTTCTATCCA 2478

Qy       581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2479 TGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACA 2538

Qy       601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2539 GTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCTGTTCAACTTCTGCAGAGA 2598

Qy       621 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2599 CTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTGGATGCAAAAGAATTGAAT 2658

Qy       641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2659 GTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAGCCTGTGCTCTTCAACTGC 2718

Qy       661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 680
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2719 AGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTGGTGCCATTAGATGAGATC 2778

Qy       681 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2779 AAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAAGGTCTCCCTGAGAAATAC 2838

Qy       701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2839 AACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCT 2898

Qy       721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2899 ATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCC 2958

Qy       741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2959 CTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGATAAAGCCAGGACGGGAAGG 3018

Qy       761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3019 ACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGGTT 3078

Qy       781 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 800
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3079 CTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAGCTCCTGAGAAATCGAGAC 3138

Qy       801 IleTyrPheLysLeuValAsnAlaGlnSerValGln 812
             ||||||||||||||||||||||||||||||||||||
Db      3139 ATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3174

RESULT 3
AX339032
LOCUS       AX339032                3621 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 14 from Patent WO0194400.
ACCESSION   AX339032
VERSION     AX339032.1  GI:18129124
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 14 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..3621
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1081 a    696 c    842 g   1002 t
ORIGIN

Alignment Scores:
Pred. No.:              2.57e-296      Length:         3621
Score:                  4079.00        Matches:        812
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            100.00%        Indels:         0
DB:                     6              Gaps:           0

US-09-873-409-2 (1-812) x AX339032 (1-3621)

Qy         1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1183 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 1242

Qy        21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1243 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 1302

Qy        41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1303 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 1362

Qy       61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1363 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 1422

Qy       81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1423 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 1482

Qy      101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1483 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 1542

Qy      121 AlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThr 140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1543 GCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTAGCACACCGACTTTCTACT 1602

Qy      141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1603 ATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATGCTGGCGGAGAAAGGAGCA 1662

Qy      161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1663 CATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTTGTGATGTCACAGGATATT 1722

Qy      181 LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer 200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1723 AAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCA 1782

Qy      201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSer 220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1783 CTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCC 1842

Qy      221 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1843 ACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAAC 1902

Qy      241 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1903 AAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCTGTTCTAAATGGAACTGTT 1962

Qy      261 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1963 CATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATGTTTGGAAATAATGATAAA 2022

Qy      281 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2023 ACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTCGTCATTTTGGGTGTTATT 2082

Qy      301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2083 TGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGAGCAGGGGAAATTTTAACG 2142

Qy      321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2143 ATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAGGATATTGCCTGGTTTGAT 2202

Qy      341 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2203 GAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCCATAGATATAGCACAAATT 2262

Qy      361 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2263 CAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAATGCAACTAACATGGGACTT 2322

Qy      381 SerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAla 400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2323 TCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTCCTGATTCTGAGTATTGCT 2382

Qy      401 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2383 CCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATGACTGGATTTGCCAACAAA 2442

Qy      421 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2443 GATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAGAATATACGT 2502

Qy      441 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2503 ACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATGTATGAAGAGATGCTTCAG 2562

Qy      461 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2563 ACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTATGCATTCAGC 2622

Qy      481 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2623 CATGCCTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTTGGAGCCTATTTAATTCAA 2682

Qy      501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2683 GCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACTGCAATTGCATATGGAGCT 2742

Qy      521 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2743 ATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCCAAAGCCAAATCGGGGGCT 2802

Qy      541 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2803 GCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGACAGCCGCAGTCAAGAAGGG 2862

Qy      561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2863 AAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTCTTCTATCCA 2922

Qy      581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2923 TGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACA 2982

Qy      601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2983 GTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCTGTTCAACTTCTGCAGAGA 3042

Qy      621 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3043 CTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTGGATGCAAAAGAATTGAAT 3102

Qy      641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3103 GTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAGCCTGTGCTCTTCAACTGC 3162

Qy      661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3163 AGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTGGTGCCATTAGATGAGATC 3222

Qy      681 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3223 AAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAAGGTCTCCCTGAGAAATAC 3282

Qy      701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3283 AACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCT 3342

Qy      721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3343 ATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCC 3402

Qy      741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3403 CTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGATAAAGCCAGGACGGGAAGG 3462

Qy      761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3463 ACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGGTT 3522
```

```
Qy     781 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3523 CTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAGCTCCTGAGAAATCGAGAC 3582

Qy     801 IleTyrPheLysLeuValAsnAlaGlnSerValGln 812
           ||||||||||||||||||||||||||||||||||||
Db    3583 ATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3618

RESULT 4
AX339031
LOCUS       AX339031                3702 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 13 from Patent WO0194400.
ACCESSION   AX339031
VERSION     AX339031.1  GI:18129123
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 13 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..3702
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1098 a    708 c    856 g   1039 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:               2.65e-296     Length:          3702
Score:                   4079.00       Matches:         812
Percent Similarity:      100.00%       Conservative:    0
Best Local Similarity:   100.00%       Mismatches:      0
Query Match:             100.00%       Indels:          0
DB:                      6             Gaps:            0

US-09-873-409-2 (1-812) x AX339031 (1-3702)

Qy       1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1264 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 1323

Qy      21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1324 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 1383

Qy      41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1384 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 1443

Qy      61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1444 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 1503

Qy      81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1504 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 1563

Qy     101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1564 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 1623

Qy     121 AlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThr 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1624 GCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTAGCACACCGACTTTCTACT 1683

Qy     141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1684 ATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATGCTGGCGGAGAAAGGAGCA 1743

Qy     161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1744 CATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTTGTGATGTCACAGGATATT 1803

Qy     181 LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1804 AAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCA 1863

Qy     201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSer 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1864 CTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCC 1923

Qy     221 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1924 ACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAAC 1983

Qy     241 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1984 AAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCTGTTCTAAATGGAACTGTT 2043

Qy     261 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2044 CATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATGTTTGGAAATAATGATAAA 2103

Qy     281 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2104 ACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTCGTCATTTTGGGTGTTATT 2163

Qy     301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2164 TGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGAGCAGGGGAAATTTTAACG 2223

Qy     321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2224 ATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAGGATATTGCCTGGTTTGAT 2283

Qy     341 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2284 GAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCCATAGATATAGCACAAATT 2343

Qy     361 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2344 CAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAATGCAACTAACATGGGACTT 2403

Qy     381 SerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAla 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2404 TCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTCCTGATTCTGAGTATTGCT 2463

Qy     401 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2464 CCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATGACTGGATTTGCCAACAAA 2523

Qy     421 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2524 GATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAGAATATACGT 2583

Qy     441 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2584 ACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATGTATGAAGAGATGCTTCAG 2643

Qy     461 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2644 ACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTATGCATTCAGC 2703

Qy     481 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2704 CATGCCTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTTGGAGCCTATTTAATTCAA 2763

Qy     501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2764 GCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACTGCAATTGCATATGGAGCT 2823
```

```
Qy     521 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2824 ATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCCAAAGCCAAATCGGGGGCT 2883

Qy     541 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2884 GCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGACAGCCGCAGTCAAGAAGGG 2943

Qy     561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2944 AAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTCTTCTATCCA 3003

Qy     581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3004 TGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACA 3063

Qy     601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3064 GTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCTGTTCAACTTCTGCAGAGA 3123

Qy     621 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3124 CTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTGGATGCAAAAGAATTGAAT 3183

Qy     641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3184 GTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAGCCTGTGCTCTTCAACTGC 3243

Qy     661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3244 AGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTGGTGCCATTAGATGAGATC 3303

Qy     681 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3304 AAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAAGGTCTCCCTGAGAAATAC 3363

Qy     701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3364 AACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCT 3423

Qy     721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3424 ATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCC 3483

Qy     741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3484 CTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGATAAAGCCAGGACGGGAAGG 3543

Qy     761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3544 ACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGGTT 3603

Qy     781 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3604 CTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAGCTCCTGAGAAATCGAGAC 3663

Qy     801 IleTyrPheLysLeuValAsnAlaGlnSerValGln 812
           ||||||||||||||||||||||||.|||||||||||
Db    3664 ATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3699
```

```
RESULT 5
AX478104
LOCUS       AX478104                3699 bp    DNA     linear   PAT 12-AUG-2002
DEFINITION  Sequence 34 from Patent WO0240541.
ACCESSION   AX478104
VERSION     AX478104.1  GI:22217064
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Tang,Y.T., Yue,H., Nguyen,D.B., Hafalia,A.J., Elliott,V.S., Lu,Y.,
            Walia,N.K., Yao,M.G., Baughn,M.R., Gandhi,A.R., Ding,L.,
            Sanjanwala,M., Ramkumar,J., Arvizu,C., Gietzen,K.J., Lal,P.G.,
            Azimzai,Y., Khan,F.A., Thangavelu,K., Thornton,M., Lu,D.A.,
            Tribouley,C.M., Warren,B.A., Ison,C.H., Das,D., Raumann,B.E.,
            Policky,J.L. and Kearney,L.
  TITLE     Transporters and ion channels
  JOURNAL   Patent: WO 0240541-A 34 23-MAY-2002;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..3699
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 7472030CB1"
BASE COUNT     1116 a    707 c    860 g   1016 t
ORIGIN

Alignment Scores:
Pred. No.:              2.62e-284       Length:         3699
Score:                  3919.00         Matches:        785
Percent Similarity:     96.80%          Conservative:   1
Best Local Similarity:  96.67%          Mismatches:     0
Query Match:            96.08%          Indels:         26
DB:                     6               Gaps:           1

US-09-873-409-2 (1-812) x AX478104 (1-3699)
```

```
Qy       1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1339 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 1398

Qy      21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1399 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 1458

Qy      41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1459 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 1518

Qy      61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1519 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 1578

Qy      81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1579 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 1638

Qy     101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1639 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 1698

Qy     121 AlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThr 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1699 GCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTAGCACACCGACTTTCTACT 1758

Qy     141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1759 ATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATGCTGGCGGAGAAAGGAGCG 1818

Qy     161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1819 CATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTTGTGATGTCACAGGATATT 1878

Qy     181 LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1879 AAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCA 1938

Qy     201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSer 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1939 CTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCC 1998

Qy     221 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    1999 ACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAAC 2058

Qy     241 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2059 AAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCTGTTCTAAATGGAACTGTT 2118

Qy     261 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2119 CATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATGTTTGGAAATAATGATAAA 2178

Qy     281 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2179 ACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTCGTCATTTTGGGTGTTATT 2238

Qy     301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 320
           |||||||||||||||||||||
Db    2239 TGCTTTGTCAGTTATTTCATG--------------------------------------- 2259

Qy     321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 340
                                                  |||||||||||||||||||||
Db    2260 ---------------------------------------CAGGATATTGCCTGGTTTGAT 2280

Qy     341 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2281 GAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCCATAGATATAGCACAAATT 2340

Qy     361 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2341 CAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAATGCAACTAACATGGGACTT 2400

Qy     381 SerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAla 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2401 TCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTCCTGATTCTGAGTATTGCT 2460

Qy     401 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2461 CCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATGACTGGATTTGCCAACAAA 2520

Qy     421 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2521 GATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAGAATATACGT 2580

Qy     441 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2581 ACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATGTATGAAGAGATGCTTCAG 2640

Qy     461 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2641 ACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTATGCATTCAGC 2700

Qy     481 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2701 CATGCCTTTATATATTTTGCCTATGCGGCAGGGTTTCGATTTGGAGCCTATTTAATTCAA 2760

Qy     501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2761 GCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACTGCAATTGCATATGGAGCT 2820

Qy     521 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540
           ||||||||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db    2821 ATGGCCATCGGAGAAACGCTCGTTTTGGCTCCTGAATATTCCAAAGCCAAATCGGGGGCT 2880

Qy     541 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2881 GCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGACAGCCGCAGTCAAGAAGGG 2940

Qy     561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2941 AAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTCTTCTATCCA 3000

Qy     581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3001 TGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACA 3060

Qy     601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3061 GTAGCATTTGTGGGGAGCAGCGGCTGTGGAAAAGCACTTCTGTTCAACTTCTGCAGAGA 3120

Qy     621 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3121 CTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTGGATGCAAAAGAATTGAAT 3180

Qy     641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3181 GTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAGCCTGTGCTCTTCAACTGC 3240

Qy     661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3241 AGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTGGTGCCATTAGATGAGATC 3300

Qy     681 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3301 AAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAAGGTCTCCCTGAGAAATAC 3360

Qy     701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3361 AACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCT 3420

Qy     721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3421 ATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCC 3480

Qy     741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3481 CTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGATAAAGCCAGGACGGGAAGG 3540

Qy     761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3541 ACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGGTT 3600

Qy     781 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3601 CTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAGCTCCTGAGAAATCGAGAC 3660

Qy     801 IleTyrPheLysLeuValAsnAlaGlnSerValGln 812
           ||||||||||||||||||||||||||||||||||||
Db    3661 ATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3696
```

```
RESULT 6
AX339027
LOCUS       AX339027               2066 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 9 from Patent WO0194400.
ACCESSION   AX339027
VERSION     AX339027.1  GI:18129119
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 9 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..2066
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      647 a    403 c    470 g    546 t
ORIGIN

Alignment Scores:
Pred. No.:              6.27e-250       Length:          2066
Score:                  3456.00         Matches:         687
Percent Similarity:     100.00%         Conservative:    0
```

```
Best Local Similarity:   100.00%      Mismatches:    0
Query Match:             84.73%       Indels:        0
DB:                      6            Gaps:          0

US-09-873-409-2 (1-812) x AX339027 (1-2066)

Qy     126 SerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAsp 145
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3 AGCAAAGGTCGGACTACAATCGTGGTAGCACACCGACTTTCTACTATTCGAAGTGCAGAT 62

Qy     146 LeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMet 165
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63 TTGATTGTGACCCTAAAGGATGGAATGCTGGCGGAGAAAGGAGCACATGCTGAACTAATG 122

Qy     166 AlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGlu 185
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 GCAAAACGAGGTCTATATTATTCACTTGTGATGTCACAGGATATTAAAAAAGCTGATGAA 182

Qy     186 GlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSer 205
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 CAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCACTTCCTCTGCACTCT 242

Qy     206 ValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGlu 225
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 GTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCCACCCAATCTAAAGAG 302

Qy     226 IleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpPro 245
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 ATAAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAACAAGCCTGAATGGCCT 362

Qy     246 PheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSer 265
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     363 TTTGTGGTTCTGGGGACATTGGCTTCTGTTCTAAATGGAACTGTTCATCCAGTATTTTCC 422

Qy     266 IleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHis 285
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     423 ATCATCTTTGCAAAAATTATAACCATGTTTGGAAATAATGATAAAACCACATTAAAGCAT 482

Qy     286 AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr 305
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     483 GATGCAGAAATTTATTCCATGATATTCGTCATTTTGGGTGTTATTTGCTTTGTCAGTTAT 542

Qy     306 PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHis 325
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     543 TTCATGCAGGGATTATTTTACGGCAGAGCAGGGGAAATTTTAACGATGAGATTAAGACAC 602

Qy     326 LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSer 345
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     603 TTGGCCTTCAAAGCCATGTTATATCAGGATATTGCCTGGTTTGATGAAAAGGAAAACAGC 662

Qy     346 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     663 ACAGGAGGCTTGACAACAATATTAGCCATAGATATAGCACAAATTCAAGGAGCAACAGGT 722

Qy     366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     723 TCCAGGATTGGCGTCTTAACACAAAATGCAACTAACATGGGACTTTCAGTTATCATTTCC 782

Qy     386 PheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     783 TTTATATATGGATGGGAGATGACATTCCTGATTCTGAGTATTGCTCCAGTACTTGCCGTG 842

Qy     406 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     843 ACAGGAATGATTGAAACCGCAGCAATGACTGGATTTGCCAACAAAGATAAGCAAGAACTT 902

Qy     426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     903 AAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAGAATATACGTACTATAGTGTCATTA 962

Qy     446 ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn 465
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     963 ACAAGGGAAAAAGCCTTCGAGCAAATGTATGAAGAGATGCTTCAGACTCAACACAGAAAT 1022

Qy     466 ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr 485
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1023 ACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTATGCATTCAGCCATGCCTTTATATAT 1082

Qy     486 PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr 505
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1083 TTTGCCTATGCAGCAGGGTTTCGATTTGGAGCCTATTTAATTCAAGCTGGACGAATGACC 1142

Qy     506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys 525
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1143 CCAGAGGGCATGTTCATAGTTTTTACTGCAATTGCATATGGAGCTATGGCCATCGGAAAA 1202

Qy     526 ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAla 545
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1203 ACGCTCGTTTTGGCTCCTGAATATTCCAAAGCCAAATCGGGGGCTGCGCATCTGTTTGCC 1262

Qy     546 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 565
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1263 TTGTTGGAAAAGAAACCAAATATAGACAGCCGCAGTCAAGAAGGGAAAAAGCCAGACACA 1322

Qy     566 CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspVal 585
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1323 TGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTCTTCTATCCATGTCGCCCAGATGTT 1382

Qy     586 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1383 TTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGCATTTGTGGGG 1442

Qy     606 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProVal 625
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1443 AGCAGCGGCTGTGGGAAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTG 1502

Qy     626 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArg 645
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1503 CAAGGACAAGTGCTGTTTGATGGTGTGGATGCAAAAGAATTGAATGTACAGTGGCTCCGT 1562

Qy     646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1563 TCCCAAATAGCAATCGTTCCTCAAGAGCCTGTGCTCTTCAACTGCAGCATTGCTGAGAAC 1622

Qy     666 IleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn 685
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1623 ATCGCCTATGGTGACAACAGCCGTGTGGTGCCATTAGATGAGATCAAAGAAGCCGCAAAT 1682

Qy     686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 705
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1683 GCAGCAAATATCCATTCTTTTATTGAAGGTCTCCCTGAGAAATACAACACACAAGTTGGA 1742

Qy     706 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1743 CTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTT 1802

Qy     726 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1803 CTCCAAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGATAATGACAGT 1862

Qy     746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1863 GAGAAGGTGGTTCAGCATGCCCTTGATAAAGCCAGGACGGGAAGGACATGCCTAGTGGTC 1922

Qy     766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1923 ACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGGTTCTGCACAATGGAAAG 1982

Qy     786 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 805
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1983 ATAAAGGAACAAGGAACTCATCAAGAGCTCCTGAGAAATCGAGACATATATTTTAAGTTA 2042

Qy     806 ValAsnAlaGlnSerValGln 812
           |||||||||||||||||||||
Db    2043 GTGAATGCACAGTCAGTGCAG 2063
```

```
RESULT 7
GGA9799
LOCUS       GGA9799              4018 bp    mRNA    linear   VRT 03-MAR-1999
DEFINITION  Gallus gallus mRNA for ABC transporter protein.
ACCESSION   AJ009799
VERSION     AJ009799.1  GI:3355756
KEYWORDS    ABC transporter protein; cmdr1 gene; P-glycoprotein.
SOURCE      Gallus gallus.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 4018)
  AUTHORS   Edelmann,H.M.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JUL-1998) Edelmann H.M.L., Institute of Molecular
            Genetics, Biocenter and University of Vienna, Dr. Bohr-Gasse 9/2
            A-1030 Vienna, A-1030 Vienna, AUSTRIA
REFERENCE   2  (bases 1 to 4018)
  AUTHORS   Edelmann,H.M., Duchek,P., Rosenthal,F.E., Foger,N., Glackin,C.,
            Kane,S.E. and Kuchler,K.
  TITLE     Cmdr1, a chicken P-glycoprotein, confers multidrug resistance and
            interacts with estradiol
  JOURNAL   Biol. Chem. 380 (2), 231-241 (1999)
  MEDLINE   99209805
   PUBMED   10195430
FEATURES             Location/Qualifiers
     source          1..4018
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /tissue_type="intestine"
     gene            109..3975
                     /gene="cmdr1"
     CDS             109..3975
                     /gene="cmdr1"
                     /note="ABC transporter protein"
                     /codon_start=1
                     /product="P-glycoprotein"
                     /protein_id="CAA08835.1"
                     /db_xref="GI:3355757"
                     /db_xref="SPTREMBL:O93437"
                     /translation="MHSEDEKQKHTVDGNTYEIANSNQDQDPEDEKKGKKKKGKKPQM
                     VSPLALFRYSSCTDKLLMIFGSLLAIAHGTSLPIAMIIFGDMTDSFVTSGMTNITGNS
                     SGLNSSADVFNKLEEEMTRYAYYYSAIAAAVLVAAYIQTSFWTLAAGRQVKKIREKFF
                     HAIMRQEIGWFDVNDAGELNTRLIDDVSKINEGIGDKIGFLIQSETTFLTGFIVGFIR
                     GWKLTLVILAVSPVLGLSAALWAKILTAFTDKEQAAYAKAGAVAEEVLSAVRTVIAFG
                     GQEKEIKRYHKNLEDAKRIGIRKAITSNISMGAAFLLIYASYALAFWYGTTLILANEY
                     SIGNVLTVFFSVLIGAFSIGQTAPSIEAFANARGAAYAIFNIIDNEPEIDSYSDAGHK
                     PDHIKGNLEFQNVFFNYPSRPDVEILKGLNLKVNCGQTVALVGGSGCGKSTTVQLIQR
                     FYDPKEGTITIDGQDLKSLNVRYLREIIGVVNQEPVLFATTIAENIRYGREDVTMEEI
                     ERATKEANAYDFIMKLPKKFETVVGERGAQMSGGQKQRIAIARALVHNPKILLLDEAT
                     SALDTESESVVQAALDKAREGRTTVVVAHRLSTVRNADLIAVFESGVITEQGNHSQLI
                     EKKGIYYKLVNMQTIETEDPSSEKSENAVSVKRSGSQSNLDESLKKELRRGSTRRSMK
                     KPGEPNDTDEKGSSPDEELPPVSFLKLMKLNKNEWPYFVAGTFCAIVNGALQPAFSVI
                     FSEIIGIFSETDQKVLREKSNLYSLLFLALGIISFFTFFVQGFAFGKAGEILTMKLRF
                     MAFKAMLRQDMAWFDDPKNSTGALTTRLANDASQVKGATGVRLALIAQNIANLGTGII
                     ISLVYGWQLTLLLLAVVPIIAVAGMIEMKMLAGHAKKDKIELEAAGKIATEAIENIRT
                     VASLTREKRFELMYGEHLLVPYRNSVKKAHIFGFCFSLSQAMMFFTYAGCFRFGAYLV
                     VNGHIEYKTVFLVFSAVVFGAMALGQTSSFAPDYAKAKISAAHLFVLFNRVPPIDSYR
                     EDGEKPEKFGGNTRIKDVKFNYPNRPEVKILQGLNLAVEKGETLALVGSSGCGKSTVV
                     QLLERFYDPLSGEIVFDDIDAKTLNIQWLRSHIGIVSQEPILFDFTIAENIAYGDNSR
                     EVSHEEIISAAKAASIHSFIDSLPEKYNTRVGDKGTQLSGGQKQRIAIARALIRKPQI
                     LLLDEATSALDTESEKIVQEALDKAREGRTCIVIAHRLSTIQNADKIAVIQNGKVIEQ
                     GTHQQLLAEKGFYYSLVNVQSGSCNM"
     mat_peptide     112..3972
                     /gene="cmdr1"
                     /product="ABC transporter protein"
                     /note="P-glycoprotein"
BASE COUNT     1288 a    758 c    911 g   1061 t
ORIGIN

Alignment Scores:
Pred. No.:              2.32e-176      Length:        4018
```

```
Score:                   2480.00      Matches:        472
Percent Similarity:      75.70%       Conservative:   151
Best Local Similarity:   57.35%       Mismatches:     186
Query Match:             60.80%       Indels:         14
DB:                      5            Gaps:           1

US-09-873-409-2 (1-812) x GGA9799 (1-4018)

Qy        2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db     1489 ATTGACGGGCAGGATCTGAAGAGCCTGAATGTAAGATATCTGCGAGAGATTATTGGTGTG 1548

Qy       22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db     1549 GTGAACCAGGAGCCCGTGCTGTTTGCTACAACTATTGCAGAAAATATTCGTTATGGCCGT 1608

Qy       42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db     1609 GAGGATGTCACCATGGAAGAAATTGAAAGAGCTACCAAGGAAGCTAATGCTTATGATTTC 1668

Qy       62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db     1669 ATCATGAAGCTCCCCAAGAAGTTTGAAACTGTGGTTGGAGAAAGAGGGGCACAGATGAGT 1728

Qy       82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db     1729 GGAGGGCAGAAGCAGCGAATAGCAATTGCCCGTGCTCTGGTTCACAATCCTAAAATTCTT 1788

Qy      102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db     1789 CTGCTTGATGAGGCAACGTCAGCTTTGGATACTGAGAGTGAATCCGTTGTGCAGGCAGCA 1848

Qy      122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db     1849 CTGGACAAGGCAAGAGAAGGACGGACCACGGTTGTAGTAGCTCATCGTCTATCAACAGTC 1908

Qy      142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db     1909 CGAAATGCAGATCTTATCGCTGTGTTTGAAAGTGGAGTTATCACAGAACAAGGAAATCAT 1968

Qy      162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLys 181
Db     1969 TCTCAACTGATTGAAAAAAAGGGAATCTACTATAAACTTGTTAACATGCAGACAATAGAA 2028

Qy      182 LysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu 201
Db     2029 ACTGAAGATCCATCATCAGAAAAAAGTGAAAATGCAGTCAGTGTCAAAAGAAGTGGATCT 2088

Qy      202 ProLeuHisSerValLysSerIleLysSerAspPhe----------------------- 213
Db     2089 CAATCAAATCTGGATGAATCCCTAAAAAAAGAATTGAGAAGAGGATCAACTAGGAGAAGC 2148

Qy      214 -----------------IleAspLysAlaGluGluSerThrGlnSerLysGluIleSer 227
Db     2149 ATGAAAAAGCCAGGAGAACCTAATGATACTGATGAGAAAGGTAGTTCACCTGATGAAGAA 2208

Qy      228 LeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheVal 247
Db     2209 TTACCTCCAGTTTCATTTCTGAAACTTATGAAGTTGAACAAAAACGAATGGCCATACTTT 2268

Qy      248 ValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIle 267
Db     2269 GTAGCTGGAACTTTCTGTGCAATTGTCAATGGAGCTTTACAACCTGCATTTTCAGTCATA 2328

Qy      268 PheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAla 287
Db     2329 TTCTCAGAAATTATAGGGATTTTTTCAGAAACTGATCAGAAAGTTTTGAGAGAAAAGAGC 2388

Qy      288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db     2389 AACTTATATTCACTGCTGTTTTTAGCACTTGGGATCATTTCCTTTTTTACTTTCTTTGTT 2448

Qy      308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
```

```
                 ||||||      :::|||:::||||||||||||||||||||:::||||||     :::|||
Db      2449 CAGGGTTTTGCATTTGGCAAAGCTGGAGAAATTCTTACAATGAAGCTTCGATTCATGGCA 2508

Qy       328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
             ||||||||||||||     ||||||:::||||||||||||:::     :::|||||||||||
Db      2509 TTTAAAGCAATGCTGAGACAGGACATGGCCTGGTTTGATGATCCTAAAAATAGCACTGGA 2568

Qy       348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
                ||||||||   ||||||   |||    :::|||::::::||||||||||||    |||
Db      2569 GCACTGACAACAAGACTTGCTAATGATGCCTCACAAGTAAAAGGAGCAACTGGTGTCAGA 2628

Qy       368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
             :::   ::::::      ||||||      |||:::|||      :::|||||||||    :::
Db      2629 TTGGCATTAATTGCCCAAAACATAGCTAACCTGGGAACTGGAATTATTATATCTTTAGTT 2688

Qy       388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
             |||||||||::::::|||   |||:::|||::::::    |||::::::||||||    |||
Db      2689 TATGGCTGGCAGCTGACCCTGTTACTCTTAGCTGTTGTACCAATCATTGCAGTGGCAGGA 2748

Qy       408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
             |||||||||      :::   |||   |||   |||||||||   ||||||:::
Db      2749 ATGATTGAAATGAAGATGTTGGCTGGGCATGCAAAAAAAGATAAAATAGAGCTGGAAGCT 2808

Qy       428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
             ||||||||||||||||||||||||:::||||||||||||||:::     ||||||||||||
Db      2809 GCAGGAAAGATTGCCACAGAAGCCATAGAAAATATTCGAACTGTTGCTAGTTTGACTCGG 2868

Qy       448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
             ||||||   ||||||   ||||||   |||   |||         :::||||||:::
Db      2869 GAAAAAAGATTTGAGCTCATGTACGGAGAACATTTGCTTGTACCATACAGAAATTCTGTG 2928

Qy       468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
             |||||||||   |||   |||   |||::::::   |||   |||   ::::::|||
Db      2929 AAGAAGGCGCACATATTTGGGTTTGTTTTCCCTTTCACAAGCAATGATGTTTTTTACC 2988

Qy       488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
             ||||||   .   ||||||||||||||||||||||||:::      |||   :::   :::
Db      2989 TATGCTGGCTGTTTCCGCTTTGGAGCCTATCTTGTGGTTAATGGACACATAGAGTACAAA 3048

Qy       508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
             :::|||:::||||||:::|||:::   :::|||||||||||||||:::|||:::|||
Db      3049 ACTGTGTTTTTAGTGTTTTCAGCTGTTGTATTTGGTGCTATGGCATTAGGACAAACCAGC 3108

Qy       528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
                   ||||||:::|||:::|||||||||      |||||||||||||||   |||
Db      3109 TCTTTTGCACCAGACTATGCCAAAGCCAAGATATCAGCAGCCCACTTGTTTGTGCTGTTT 3168

Qy       548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
                :::   |||   |||||||||      ::::::|||:::||||||:::
Db      3169 AATAGAGTACCTCCAATAGACAGTTACCGCGAGGATGGAGAGAAACCTGAAAAATTTGGG 3228

Qy       568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
             ||||||         ::::::|||   |||   ||||||   ||||||:::|||   |||
Db      3229 GGAAACACAAGGATCAAAGATGTGAAATTTAACTATCCAAACCGACCAGAGGTCAAAATC 3288

Qy       588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
             |||:::||||||:::||||||::::::|||:::|||:::|||:::   ||||||||||||
Db      3289 CTCCAAGGCTTGAATCTAGCAGTAGAAAAAGGAGAAACGTTGGCCCTTGTTGGTAGCAGT 3348

Qy       608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
             ||||||||||||||||||   ||||||||||||:::|||   |||||||||:::    |||
Db      3349 GGATGTGGAAAGAGCACTGTTGTTCAGCTACTTGAGAGATTTTATGATCCACTCAGTGGA 3408

Qy       628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
             :::::::::||||||   :::|||||||||   ||||||:::||||||||||||||||||
Db      3409 GAAATTGTGTTTGATGACATTGATGCAAAGACACTAAATATCCAGTGGCTGAGATCTCAC 3468

Qy       648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
             |||   ||||||   |||||||||:::||||||:::    :::|||||||||||||||||||
Db      3469 ATCGGTATCGTCTCACAAGAGCCAATCCTGTTCGACTTCACCATCGCTGAAAACATCGCG 3528

Qy       668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
             ||||||||||||||||||   |||      :::||||||      ||||||   ||||||

Db      3529 TACGGGGACAACAGTCGGGAGGTGTCCCATGAGGAAATCATTAGTGCAGCGAAAGCAGCC 3588

Qy       688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
             :::|||||||||||||||:::   |||||||||||||||||||||:::||||||   |||
Db      3589 AGTATTCATTCCTTCATTGACTCTCTGCCAGAAAAATACAATACTCGTGTAGGAGACAAG 3648

Qy       708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
             |||   ||||||||||||||||||:::||||||||||||||||||||||||::::::
Db      3649 GGAACACAGCTTTCTGGTGGTCAGAAACAACGTATTGCTATTGCTCGAGCTCTTATACGT 3708

Qy       728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
             ||||||:::|||||||||||||||||||||||||||||||||||||    :::|||||||||
Db      3709 AAGCCCCAGATTCTGCTACTGGATGAAGCTACATCTGCCTTGGACACAGAAAGCGAAAAG 3768

Qy       748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
             :::||||||   ||||||||||||||||||   ||||||||||||:::|||:::    |||
Db      3769 ATTGTCCAGGAAGCACTGGATAAAGCCAGAGAAGGTCGCACCTGCATCGTGATAGCTCAC 3828

Qy       768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
             |||||||||   ||||||||||||||||||   |||   |||:::   |||||||||:::
Db      3829 CGCCTGTCCACCATCCAGAATGCTGACAAGATTGCTGTGATCCAGAACGGCAAGGTCATA 3888

Qy       788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
             ||||||||||||||||||:::||||||         :::      |||:::   |||||||||
Db      3889 GAGCAAGGGACTCACCAGCAGCTCCTGGCTGAAAAAGGCTTCTACTACTCGCTGGTCAAT 3948

Qy       808 AlaGlnSer 810
                ||||||
Db      3949 GTTCAAAGT 3957
```

```
RESULT 8
CRUPGP1185
LOCUS       CRUPGP1185              3987 bp    mRNA    linear   ROD 27-APR-1993
DEFINITION  Chinese hamster p-glycoprotein mRNA (clone ADX185), complete cds.
ACCESSION   M59254
VERSION     M59254.1  GI:191156
KEYWORDS    multidrug resistance glycoprotein; p-glycoprotein.
SOURCE      Chinese hamster, cDNA to mRNA, clone ADX185.
  ORGANISM  Cricetulus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
            Cricetulus.
REFERENCE   1  (bases 1 to 3987)
  AUTHORS   Devine,S.E., Hussain,A., Davide,J.P. and Melera,P.W.
  TITLE     Full length and alternatively spliced pgp1 transcripts in
            multidrug-resistant Chinese hamster lung cells
  JOURNAL   J. Biol. Chem. 266 (7), 4545-4555 (1991)
  MEDLINE   91154265
   PUBMED   1671863
FEATURES             Location/Qualifiers
     source          1..3987
                     /organism="Cricetulus sp."
                     /db_xref="taxon:10031"
                     /cell_type="fibroblast"
                     /tissue_type="lung"
     gene            1..3987
                     /gene="pgp-1"
     CDS             115..3624
                     /gene="pgp-1"
                     /codon_start=1
                     /product="p-glycoprotein"
                     /protein_id="AAA37005.1"
                     /db_xref="GI:191157"
                     /translation="MTTYAYYYTGIGAGVLIVAYIQVSFWCLAAGRQIHKIRQKFFHA
                     IMNQEIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQAMATFFGGFIIGFTRGW
                     KLTLVILAISPVLGLSAGIWAKILSSFTDKELQAYAKAGAVAEEVLAAIRTVIAFGGQ
                     KKELERYNNNLEEAKRLGIKKAITANISMGAAFLLIYASYALAFWYGTSLVISKEYSI
                     GQVLTVFFAVLIAPFSIGQASPNIEAFANARGAAYEIFNIIDNKPSIDSFSKNGYKPD
                     NIKGNLEFKNIHFSYPSRKDVQILKGLNLKVQSGQTVALVGNSGCGKSTTVQLLQRLY
                     DPTEGVVSIDGQDIRTINVRYLREIIGVVSQEPVLFATTIAENIRYGRENVTMDEIEK
                     AVKEANAYDFIMKLPHKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSA
                     LDTESEAVVQAALDKAREGRTTIVIAHRLSTVRNADIIAGFDGGVIVEQGNHEELMRE
                     KGIYFKLVMTQTAGNEIELGNEVGESKNEIDNLDMSSKDSASSLIRRRSTRRSIRGPH
```

```
                    DQDRKLSTKEALDEDVPPISFWRILKLNSSEWPYFVVGIFCAIVNGALQPAFSIIFSK
                    VVGVFTRNTDDETKRHDSNLFSLLFLILGVISFITFFLQGFTFGKAGEILTKRLRYMV
                    FKSMLRQDVSWFDNPKNTTGALTTRLANDAGQVKGATGARLAVITQNIANLGTGIIIS
                    LIYGWQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGSGKIATEAIENFRTVV
                    SLTREQKFENMYAQSLQIPYRNALKKAHVFGITFSFTQAMMYFSYAACFRFGAYLVAR
                    ELMTFENVLLVFSAIVFGAMAVGQVSSFAPDYAKAKVSASHIIMIIEKVPSIDSYSTG
                    GLKPNTLEGNVKFNEVVFNYPTRPDIPVLQGLNLEVKKGQTLALVGSSGCGKSTVVQL
                    LERFYDPMAGTVFLDGKEVNQLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVV
                    SQDEIERAAKEANIHQFIESLPDKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILL
                    LDEATSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVIQNGKVKEHGT
                    HQQLLAQKGIYFSMVSVQAGAKR"
BASE COUNT     1194 a    822 c    944 g   1027 t
ORIGIN      Map position 1q26.

Alignment Scores:
Pred. No.:               4.58e-176      Length:         3987
Score:                   2476.00        Matches:        468
Percent Similarity:      77.52%         Conservative:   170
Best Local Similarity:   56.87%         Mismatches:     171
Query Match:             60.70%         Indels:         14
DB:                      10             Gaps:           4

US-09-873-409-2 (1-812) x CRUPGP1185 (1-3987)

Qy        2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
              :::|||     |||||||||    :::||||||||:::    |||:::    ||||||||
Db     1141 ATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTC 1200

Qy       22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
            ||||||||||||||||||||||||     |||||||||:::   ||||||:::||||||||
Db     1201 GTGAGTCAGGAACCTGTGTTGTTTGCCACCACTATCGCTGAAAACATTCGCTATGGCCGA 1260

Qy       42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
            ::::::||||||   :::|||:::|||:::|||   :::|||||||||||||||||||||
Db     1261 GAAAATGTCACTATGGACGAGATCGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTC 1320

Qy       62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
            ||||||:::   |||:::||||||:::||||||||||||:::|||||||||:::||||||
Db     1321 ATCATGAAACTGCCCCATAAATTTGACACTCTTGTTGGTGAGAGAGGAGCACAGCTGAGT 1380

Qy       82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1381 GGAGGACAGAAACAGAGAATCGCCATTGCTCGTGCCCTGGTCCGCAACCCCAAGATCCTT 1440

Qy      102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
            :::|||||||||||||||||||||||||||:::||||||::::::   ||||||||||||
Db     1441 TTGTTGGATGAGGCGACATCAGCCTTGGACACAGAAAGTGAAGCCGTGGTTCAGGCTGCT 1500

Qy      122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
            |||:::||||||   :::||||||||||||||||||:::||||||||||||||||||:::
Db     1501 CTGGATAAGGCTAGAGAAGGCCGGACTACCATTGTGATAGCCCACCGCTTGTCTACAGTT 1560

Qy      142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
            |||:::||||||:::|||               |||::::::   |||:::|||   |||
Db     1561 CGAAATGCTGACATCATAGCTGGGTTTGATGGTGGTGTCATTGTGGAGCAAGGAAATCAT 1620

Qy      162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178
               |||||||||   ::::::|||:::|||   |||||||||:::|||
Db     1621 GAAGAGCTCATGAGAGAGAAGGGCATTTACTTCAAACTTGTCATGACACAGACAGCAGGA 1680

Qy      179 ---------------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 191
                                 ::::::   ::::::   :::::::::::::::
Db     1681 AATGAAATTGAATTAGGAAATGAAGTTGGTGAGTCTAAAAATGAAATTGATAATTTAGAC 1740

Qy      192 TyrSerThrGluArgLysThrAsnSerLeu------ProLeuHisSerValLysSerIle 209
               |||::::::         :::||||||             :::   :::||||||
Db     1741 ATGTCTTCAAAAGATTCAGCATCCAGTCTAATCAGAAGGAGATCAACTCGCAGAAGTATC 1800

Qy      210 LysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLysGluIleSerLeu 228
            :::                   :::   ||||||   ::::::   :::         :::
Db     1801 CGTGGACCACATGACCAAGACAGGAAGCTTAGTACCAAAGAGGCCTTGGATGAAGATGTA 1860

Qy      229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
            |||   :::|||         :::||||||||||||||     |||||||||:::   |||
Db     1861 CCTCCAATTTCCTTTTGGCGTATCCTAAAGTTGAATTCATCTGAATGGCCATATTTTGTG 1920

Qy      249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
            :::|||         :::::::::||||||   :::      |||   |||||||||||||
Db     1921 GTTGGTATATTCTGTGCCATAGTAAATGGAGCCTTGCAACCAGCATTCTCAATAATATTT 1980

Qy      269 AlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAla 287
            :::|||::::::   :::|||            |||   |||         :::||||||:::
Db     1981 TCTAAGGTTGTAGGGGTTTTCACAAGAAATACTGATGATGAAACCAAACGACATGATAGC 2040

Qy      288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
               ::::::|||::::::|||:::||||||||||||||||   |||:::::::::|||:::
Db     2041 AACTTGTTTTCCTTATTGTTTCTGATCCTTGGGGTCATTTCTTTTATTACATTTTTTCTT 2100

Qy      308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
               ||||||         :::|||:::||||||||||||||||||   ||||||||::::::
Db     2101 CAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGACTCCGATACATGGTT 2160

Qy      328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
            ||||||:::||||||   ||||||::::::||||||||         :::|||:::||||||
Db     2161 TTCAAATCCATGCTGAGACAGGATGTGAGCTGGTTTGATAACCCTAAAAACACCACTGGA 2220

Qy      348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
               ||||||   ||||||      |||      |||:::||||||||||||||||:::::
Db     2221 GCATTGACCACCAGGCTCGCCAACGATGCTGGTCAAGTTAAAGGGGCTACAGGAGCCAGA 2280

Qy      368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
            :::      |||:::|||||||||         |||:::|||      :::||||||||   |||
Db     2281 CTTGCTGTCATTACCCAGAATATAGCAAATCTTGGGACAGGAATCATCATATCCCTAATC 2340

Qy      388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
            |||||||||::::::|||   |||:::|||:::|||   |||::::::|||:::   |||
Db     2341 TATGGCTGGCAGTTGACACTTCTACTCTTAGCAATTGTCCCCATCATTGCAATAGCAGGA 2400

Qy      408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
            ::::::|||         ::::::|||   |||   |||||||||:::||||||:::
Db     2401 GTGGTTGAGATGAAAATGTTGTCTGGACAAGCACTAAAAGATAAGAAGGAGCTAGAAGGT 2460

Qy      428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
            :::||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db     2461 TCTGGGAAGATTGCTACTGAAGCAATAGAGAACTTCCGCACTGTCGTCTCTTTGACTCGG 2520

Qy      448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
            |||:::   ||||||   ||||||   :::   ||||||      :::||||||
Db     2521 GAGCAGAAGTTTGAAAATATGTATGCCCAGAGCTTGCAGATACCATACAGAAATGCTCTG 2580

Qy      468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
            ||||||||   :::   |||      ::::::|||:::   |||   :::||||||:::
Db     2581 AAGAAAGCACACGTCTTTGGGATCACATTCTCCTTCACCCAGGCCATGATGTATTTTTCG 2640

Qy      488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
            ||||||||   ||||||||||||||||||||||||:::            ||||||   |||
Db     2641 TATGCTGCTTGTTTCCGGTTTGGTGCTTACTTGGTGGCACGGGAATTAATGACATTTGAA 2700

Qy      508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
               :::   :::|||||||||:::||||||   :::||||||||||||:::|||:::
Db     2701 AATGTTCTATTAGTATTCTCAGCTATTGTCTTTGGTGCCATGGCAGTTGGGCAGGTCAGT 2760

Qy      528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
                  ||||||:::|||:::||||||||      ||||:::|||:::         ::::::
Db     2761 TCATTTGCTCCTGACTATGCCAAAGCCAAAGTGTCGGCATCGCATATCATCATGATCATT 2820

Qy      548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
            ||||||   ||||:::||||||||   |||         |||   ||||||:::|||   |||
Db     2821 GAAAAAGTCCCTTCCATTGACAGCTACAGCACCGGAGGCCTGAAGCCTAATACATTGGAA 2880

Qy      568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
            ||||||::::::|||   ||||||   |||   ||||||   |||||||||:::   :::
Db     2881 GGAAATGTGAAATTTAATGAAGTCGTGTTCAACTACCCCACCCGGCCTGACATCCCAGTG 2940

Qy      588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
            |||:::||||||:::|||   :::::::::|||:::|||:::|||   ||||||||||||
```

```
Db    2941 CTTCAGGGGCTGAACCTGGAGGTGAAGAAGGGCCAGACGCTGGCCCTGGTGGGCAGCAGT 3000

Qy     608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
           ||||||||||||||||||  |||||||||||:::|||   |||||||||:::   |||
Db    3001 GGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAGCGGTTCTACGACCCCATGGCCGGA 3060

Qy     628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
           |||      ||||||   :::       :::||||||||||||||||||||||:::
Db    3061 ACAGTGTTTCTAGATGGCAAAGAAGTAAACCAACTAAATGTCCAGTGGCTCCGAGCACAC 3120

Qy     648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
           :::   ||||||   |||||||||:::||||||:::||||||||||||||||||||||||
Db    3121 CTAGGCATTGTGTCCCAAGAGCCCATCCTGTTTGACTGCAGCATCGCTGAGAACATCGCT 3180

Qy     668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
           |||||||||||||||||||||||||       ||||||||||:::   ||||||     |||
Db    3181 TATGGAGACAACAGCCGGGTCGTGTCCCAGGATGAGATCGAGAGGGCGGCGAAGGAAGCC 3240

Qy     688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
           ||||||||   |||||||||   ||||||:::||||||||||||:::||||||   |||
Db    3241 AACATCCACCAGTTCATTGAGTCACTGCCTGACAAATACAACACCAGAGTAGGAGACAAA 3300

Qy     708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
           |||   |||||||||||||||||||||||||:::||||||||||||||||||:::::
Db    3301 GGGACTCAGCTGTCGGGTGGGCAGAAACAACGCATCGCCATAGCACGCGCCCTCGTCAGA 3360

Qy     728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
           :::|||   ||||||||||||||||||||||||||||||||||    :::||||||||
Db    3361 CAGCCTCACATTTTGCTCCTGGATGAAGCAACATCAGCCCTGGATACAGAGAGTGAAAAG 3420

Qy     748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
           ||||||||   ||||||||||||||||||||    ||||||||||:::|||:::   |||
Db    3421 GTTGTCCAGGAAGCTCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATCGCTCAC 3480

Qy     768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
           ||||||||   |||||||||||||||||||||||||:::   |||||||||:::|||
Db    3481 CGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATTCAGAATGGCAAGGTCAAG 3540

Qy     788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
           |||   |||||||||||:::||||||          :::   ||||||||    :::|||:::
Db    3541 GAGCATGGCACCCACCAGCAGCTGCTGGCACAGAAAGGCATCTATTTCTCCATGGTCAGT 3600

Qy     808 AlaGlnSer 810
              |||:::
Db    3601 GTGCAGGCT 3609
```

RESULT 9
CRUPGPI

```
LOCUS       CRUPGPI                 4296 bp    mRNA    linear   ROD 22-JUN-1995
DEFINITION  C.griseus P-glycoprotein (isoform class I) mRNA, complete cds.
ACCESSION   M60040
VERSION     M60040.1  GI:191164
KEYWORDS    multidrug resistance; p-glycoprotein; transmembrane protein.
SOURCE      C.griseus adult liver and ovary, cDNA to mRNA.
  ORGANISM  Cricetulus griseus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
            Cricetulus.
REFERENCE   1  (bases 1 to 4296)
  AUTHORS   Endicott,J.A., Sarangi,F. and Ling,V.
  TITLE     Complete cDNA sequences encoding the Chinese hamster P-glycoprotein
            gene family
  JOURNAL   DNA Seq. 2 (2), 89-101 (1991)
  MEDLINE   92135896
   PUBMED   1685679
FEATURES             Location/Qualifiers
     source          1..4296
                     /organism="Cricetulus griseus"
                     /db_xref="taxon:10029"
                     /tissue_type="liver and ovary"
                     /dev_stage="adult"
     CDS             103..3933
                     /note="putative"
                     /codon_start=1
                     /product="p-glycoprotein isoform I"
                     /protein_id="AAA68883.1"
                     /db_xref="GI:191165"
                     /translation="MEFEEDFSGRKDKNFLKMGRKSKKEKKEKKPVVSVFTMFRYAGW
                     LDRLYMLVGTLAAIIHGVALPLMMLVFGDMTDSFASVGNIPTNATNNATQVNASDIFG
                     KLEEEMTTYAYYYTGIGAGVLIVAYIQVSFWCLAAGRQIHKIRQKFFHAIMNQEIGWF
                     DVHDVGELNTRLTDDVSKINEGIGDKIGMFFQAMATFFGGFIIGFTRGWKLTLVILAI
                     SPVLGLSAGIWAKILSSFTDKELQAYAKAGAVAEEVLAAIRTVIAFGGQKKELERYNN
                     NLEEAKRLGIKKAITANISMGAAFLLIYASYALAFWYGTSLVISKEYSIGQVLTVFFA
                     VLIGAFSIGQASPNIEAFANARGAAYEIFNIIDNKPSIDSFSKNGYKPDNIKGNLEFK
                     NIHFSYPSRKDVQILKGLNLKVQSGQTVALVGNSGCGKSTTVQLLQRLYDPTEGVVSI
                     DGQDIRTINVRYLREIIGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKEANAYD
                     FIMKLPHKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTESEAVV
                     QAALDKAREGRTTIVIAHRLSTVRNADIIAGFDGGVIVEQGNHEELMREKGIYFKLVM
                     TQTAGNEIELGNEVGESKNEIDNLDMSSKDSASSLIRRRSTRRSIRGPHDQDRKLSTK
                     EALDEDVPPISFWRILKLNSSEWPYFVVGIFCAIVNGALQPAFSIIFSKVVGVFTRNT
                     DDETKRHDSNLFSLLFLILGVISFITFFLQGFTFGKAGEILTKRLRYMVFKSMLRQDV
                     SWFDNPKNTTGALTTRLANDAGQVKGATGARLAVITQNIANLGTGIIISLIYGWQLTL
                     LLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGSGKIATEAIENFRTVVSLTREQKFE
                     NMYAQSLQIPYRNALKKAHVFGITFSFTQAMMYFSYAACFRFGAYLVARELMTFENVL
                     LVFSAIVFGAMAVGQVSSFAPDYAKAKVSASHIIMIIEKVPSIDSYSTGGLKPNTLEG
                     NVKFNEVVFNYPTRPDIPVLQGLNLEVKKGQTLALVGSSGCGKSTVVQLLERFYDPMA
                     GTVFLDGKEVNQLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSQDEIERAA
                     KEANIHQFIESLPDKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEATSALD
                     TESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVIQNGKVKEHGTHQQLLAQKG
                     IYFSMVSVQAGAKR"
BASE COUNT     1271 a    892 c   1029 g   1104 t
ORIGIN
```

```
Alignment Scores:
Pred. No.:              5.06e-176       Length:         4296
Score:                  2476.00         Matches:        468
Percent Similarity:     77.52%          Conservative:   170
Best Local Similarity:  56.87%          Mismatches:     171
Query Match:            60.70%          Indels:         14
DB:                     10              Gaps:           4
```

US-09-873-409-2 (1-812) x CRUPGPI (1-4296)

```
Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
           :::|||      ||||||||    :::||||||||:::   |||:::   |||||||||
Db    1450 ATCGACGGTCAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTC 1509

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           ||||||||||||||||||||||||   ||||||||:::   ||||||:::|||||||||
Db    1510 GTGAGTCAGGAACCTGTGTTGTTTGCCACCACTATCGCTGAAAACATTCGCTATGGCCGA 1569

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
           ::::::||||||   :::|||:::|||:::|||   :::||||||||||||||||||||||
Db    1570 GAAAATGTCACTATGGACGAGATCGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTC 1629

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           ||||||:::   |||:::||||||:::||||||||||||||||:::||||||||||:::|||
Db    1630 ATCATGAAACTGCCCCATAAATTTGACACTCTTGTTGGTGAGAGAGGAGCACAGCTGAGT 1689

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1690 GGAGGACAGAAACAGAGAATCGCCATTGCTCGTGCCCTGGTCCGCAACCCCAAGATCCTT 1749

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
           :::||||||||||||||||||||||||:::||||||::::::   ||||||||||||
Db    1750 TTGTTGGATGAGGCGACATCAGCCTTGGACACAGAAAGTGAAGCCGTGGTTCAGGCTGCT 1809

Qy     122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
           |||:::||||||   :::|||||||||||||||||||||:::|||||||||||||||||:::
Db    1810 CTGGATAAGGCTAGAGAAGGCCGGACTACCATTGTGATAGCCCACCGCTTGTCTACAGTT 1869

Qy     142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
           |||:::|||||||:::|||                |||::::::   |||:::|||   |||
Db    1870 CGAAATGCTGACATCATAGCTGGGTTTGATGGTGGTGTCATTGTGGAGCAAGGAAATCAT 1929

Qy     162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178
```

```
Db   1930 GAAGAGCTCATGAGAGAGAAGGGCATTTACTTCAAACTTGTCATGACACAGACAGCAGGA 1989

Qy    179 --------------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 191

Db   1990 AATGAAATTGAATTAGGAAATGAAGTTGGTGAGTCTAAAAATGAAATTGATAATTTAGAC 2049

Qy    192 TyrSerThrGluArgLysThrAsnSerLeu------ProLeuHisSerValLysSerIle 209

Db   2050 ATGTCTTCAAAAGATTCAGCATCCAGTCTAATCAGAAGGAGATCAACTCGCAGAAGTATC 2109

Qy    210 LysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLysGluIleSerLeu 228

Db   2110 CGTGGACCACATGACCAAGACAGGAAGCTTAGTACCAAAGAGGCCTTGGATGAAGATGTA 2169

Qy    229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248

Db   2170 CCTCCAATTTCCTTTTGGCGTATCCTAAAGTTGAATTCATCTGAATGGCCATATTTTGTG 2229

Qy    249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268

Db   2230 GTTGGTATATTCTGTGCCATAGTAAATGGAGCCTTGCAACCAGCATTCTCAATAATATTT 2289

Qy    269 AlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAla 287

Db   2290 TCTAAGGTTGTAGGGGTTTTCACAAGAAATACTGATGATGAAACCAAACGACATGATAGC 2349

Qy    288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307

Db   2350 AACTTGTTTTCCTTATTGTTTCTGATCCTTGGGGTCATTTCTTTTATTACATTTTTTCTT 2409

Qy    308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327

Db   2410 CAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGACTCCGATACATGGTT 2469

Qy    328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347

Db   2470 TTCAAATCCATGCTGAGACAGGATGTGAGCTGGTTTGATAACCCTAAAAACACCACTGGA 2529

Qy    348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367

Db   2530 GCATTGACCACCAGGCTCGCCAACGATGCTGGTCAAGTTAAAGGGGCTACAGGAGCCAGA 2589

Qy    368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387

Db   2590 CTTGCTGTCATTACCCAGAATATAGCAAATCTTGGGACAGGAATCATCATATCCCTAATC 2649

Qy    388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407

Db   2650 TATGGCTGGCAGTTGACACTTCTACTCTTAGCAATTGTCCCCATCATTGCAATAGCAGGA 2709

Qy    408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427

Db   2710 GTGGTTGAGATGAAAATGTTGTCTGGACAAGCACTAAAAGATAAGAAGGAGCTAGAAGGT 2769

Qy    428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447

Db   2770 TCTGGGAAGATTGCTACTGAAGCAATAGAGAACTTCCGCACTGTCGTCTCTTTGACTCGG 2829

Qy    448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467

Db   2830 GAGCAGAAGTTTGAAAATATGTATGCCCAGAGCTTGCAGATACCATACAGAAATGCTCTG 2889

Qy    468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487

Db   2890 AAGAAAGCACACGTCTTTGGGATCACATTCTCCTTCACCCAGGCCATGATGTATTTTTCG 2949

Qy    488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507

Db   2950 TATGCTGCTTGTTTCCGGTTTGGTGCTTACTTGGTGGCACGGGAATTAATGACATTTGAA 3009

Qy    508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527

Db   3010 AATGTTCTATTAGTATTCTCAGCTATTGTCTTTGGTGCCATGGCAGTTGGGCAGGTCAGT 3069

Qy    528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547

Db   3070 TCATTTGCTCCTGACTATGCCAAAGCCAAAGTGTCGGCATCGCATATCATCATGATCATT 3129

Qy    548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567

Db   3130 GAAAAAGTCCCTTCCATTGACAGCTACAGCACCGGAGGCCTGAAGCCTAATACATTGGAA 3189

Qy    568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587

Db   3190 GGAAATGTGAAATTTAATGAAGTCGTGTTCAACTACCCCACCCGGCCTGACATCCCAGTG 3249

Qy    588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607

Db   3250 CTTCAGGGGCTGAACCTGGAGGTGAAGAAGGGCCAGACGCTGGCCCTGGTGGGCAGCAGT 3309

Qy    608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627

Db   3310 GGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAGCGGTTCTACGACCCCATGGCCGGA 3369

Qy    628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647

Db   3370 ACAGTGTTTCTAGATGGCAAAGAAGTAAACCAACTAAATGTCCAGTGGCTCCGAGCACAC 3429

Qy    648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667

Db   3430 CTAGGCATTGTGTCCCAAGAGCCCATCCTGTTTGACTGCAGCATCGCTGAGAACATCGCT 3489

Qy    668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687

Db   3490 TATGGAGACAACAGCCGGGTCGTGTCCCAGGATGAGATCGAGAGGGCGGCGAAGGAAGCC 3549

Qy    688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707

Db   3550 AACATCCACCAGTTCATTGAGTCACTGCCTGACAAATACAACACCAGAGTAGGAGACAAA 3609

Qy    708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727

Db   3610 GGGACTCAGCTGTCGGGTGGGCAGAAACAACGCATCGCCATAGCACGCGCCCTCGTCAGA 3669

Qy    728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747

Db   3670 CAGCCTCACATTTTGCTCCTGGATGAAGCAACATCAGCCCTGGATACAGAGAGTGAAAAG 3729

Qy    748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767

Db   3730 GTTGTCCAGGAAGCTCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATCGCTCAC 3789

Qy    768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787

Db   3790 CGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATTCAGAATGGCAAGGTCAAG 3849

Qy    788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807

Db   3850 GAGCATGGCACCCACCAGCAGCTGCTGGCACAGAAAGGCATCTATTTCTCCATGGTCAGT 3909

Qy    808 AlaGlnSer 810

Db   3910 GTGCAGGCT 3918
```

```
RESULT 10
CRUPGP1165
LOCUS       CRUPGP1165         4304 bp    mRNA     linear   ROD 27-APR-1993
DEFINITION  Chinese hamster p-glycoprotein mRNA (clone ADX165), complete cds.
ACCESSION   M59253
VERSION     M59253.1  GI:191154
KEYWORDS    multidrug resistance glycoprotein; p-glycoprotein.
SOURCE      Chinese hamster, cDNA to mRNA, clone ADX165.
  ORGANISM  Cricetulus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
```

```
            Cricetulus.
REFERENCE   1  (bases 1 to 4304)
  AUTHORS   Devine,S.E., Hussain,A., Davide,J.P. and Melera,P.W.
  TITLE     Full length and alternatively spliced pgp1 transcripts in
            multidrug-resistant Chinese hamster lung cells
  JOURNAL   J. Biol. Chem. 266 (7), 4545-4555 (1991)
  MEDLINE   91154265
   PUBMED   1671863
FEATURES             Location/Qualifiers
     source          1..4304
                     /organism="Cricetulus sp."
                     /db_xref="taxon:10031"
                     /cell_type="fibroblast"
                     /tissue_type="lung"
     gene            1..4304
                     /gene="pgp-1"
     CDS             111..3941
                     /gene="pgp-1"
                     /codon_start=1
                     /product="p-glycoprotein"
                     /protein_id="AAA37004.1"
                     /db_xref="GI:191155"
                     /translation="MEFEEDFSGRKDKNFLKMGRKSKKEKKEKKPVVSVFTMFRYAGW
                     LDRLYMLVGTLAAIIHGVALPLMMLVFGDMTDSFASVGNIPTNATNNATQVNASDIFG
                     KLEEEMTTYAYYYTGIGAGVLIVAYIQVSFWCLAAGRQIHKIRQKFFHAIMNQEIGWF
                     DVHDVGELNTRLTDDVSKINEGIGDKIGMFFQAMATFFGGFIIGFTRGWKLTLVILAI
                     SPVLGLSAGIWAKILSSFTDKELQAYAKAGAVAEEVLAAIRTVIAFGGQKKELERYNN
                     NLEEAKRLGIKKAITANISMGAAFLLIYASYALAFWYGTSLVISKEYSIGQVLTVFFA
                     VLIAPFSIGQASPNIEAFANARGAAYEIFNIIDNKPSIDSFSKNGYKPDNIKGNLEFK
                     NIHFSYPSRKDVQILKGLNLKVQSGQTVALVGNSGCGKSTTVQLLQRLYDPTEGVVSI
                     DGQDIRTINVRYLREIIGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKEANAYD
                     FIMKLPHKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTESEAVV
                     QAALDKAREGRTTIVIAHRLSTVRNADIIAGFDGGVIVEQGNHEELMREKGIYFKLVM
                     TQTAGNEIELGNEVGESKNEIDNLDMSSKDSASSLIRRRSTRRSIRGPHDQDRKLSTK
                     EALDEDVPPISFWRILKLNSSEWPYFVVGIFCAIVNGALQPAFSIIFSKVVGVFTRNT
                     DDETKRHDSNLFSLLFLILGVISFITFFLQGFTFGKAGEILTKRLRYMVFKSMLRQDV
                     SWFDNPKNTTGALTTRLANDAGQVKGATGARLAVITQNIANLGTGIIISLIYGWQLTL
                     LLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGSGKIATEAIENFRTVVSLTREQKFE
                     NMYAQSLQIPYRNALKKAHVFGITFSFTQAMMYFSYAACFRFGAYLVARELMTFENVL
                     LVFSAIVFGAMAVGQVSSFAPDYAKAKVSASHIIMIIEKVPSIDSYSTGGLKPNTLEG
                     NVKFNEVVFNYPTRPDIPVLQGLNLEVKKGQTLALVGSSGCGKSTVVQLLERFYDPMA
                     GTVFLDGKEVNQLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSQDEIERAA
                     KEANIHQFIESLPDKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEATSALD
                     TESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVIQNGKVKEHGTHQQLLAQKG
                     IYFSMVSVQAGAKR"
BASE COUNT     1275 a    895 c   1031 g   1103 t
ORIGIN      Map position 1q26.

Alignment Scores:
Pred. No.:              5.07e-176       Length:          4304
Score:                  2476.00         Matches:         468
Percent Similarity:     77.52%          Conservative:    170
Best Local Similarity:  56.87%          Mismatches:      171
Query Match:            60.70%          Indels:          14
DB:                     10              Gaps:            4

US-09-873-409-2 (1-812) x CRUPGP1165 (1-4304)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
          :::|||      ||||||||   :::|||||||||:::   |||:::   ||||||||
Db   1458 ATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTC 1517

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
          ||||||||||||||||||||||||||||||:::   ||||||||:::||||||||||||
Db   1518 GTGAGTCAGGAACCTGTGTTGTTTGCCACCACTATCGCTGAAAACATTCGCTATGGCCGA 1577

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
          ::::::||||||   :::|||:::|||:::|||   :::||||||||||||||||||||
Db   1578 GAAAATGTCACTATGGACGAGATCGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTC 1637

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
          ||||||:::   |||:::||||||:::||||||||||||||:::|||||||||:::|||
Db   1638 ATCATGAAACTGCCCCATAAATTTGACACTCTTGTTGGTGAGAGAGGAGCACAGCTGAGT 1697

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1698 GGAGGACAGAAACAGAGAATCGCCATTGCTCGTGCCCTGGTCCGCAACCCCAAGATCCTT 1757

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
          :::|||||||||||||||||||||||||||:::||||||::::::   |||||||||||
Db   1758 TTGTTGGATGAGGCGACATCAGCCTTGGACACAGAAAGTGAAGCCGTGGTTCAGGCTGCT 1817

Qy    122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
          |||:::||||||   :::|||||||||||||||||||:::||||||||||||||||||:::
Db   1818 CTGGATAAGGCTAGAGAAGGCCGGACTACCATTGTGATAGCCCACCGCTTGTCTACAGTT 1877

Qy    142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
          |||:::||||||:::|||            |||::::::   |||:::|||   |||
Db   1878 CGAAATGCTGACATCATAGCTGGGTTTGATGGTGGTGTCATTGTGGAGCAAGGAAATCAT 1937

Qy    162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178
             ||||||||   ::::::|||:::|||:::   |||||||||:::|||
Db   1938 GAAGAGCTCATGAGAGAGAAGGGCATTTACTTCAAACTTGTCATGACACAGACAGCAGGA 1997

Qy    179 --------------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 191
                              ::::::   ::::::      ::::::::::::::
Db   1998 AATGAAATTGAATTAGGAAATGAAGTTGGTGAGTCTAAAAATGAAATTGATAATTTAGAC 2057

Qy    192 TyrSerThrGluArgLysThrAsnSerLeu------ProLeuHisSerValLysSerIle 209
             |||::::::            :::||||||             :::   :::||||||
Db   2058 ATGTCTTCAAAAGATTCAGCATCCAGTCTAATCAGAAGGAGATCAACTCGCAGAAGTATC 2117

Qy    210 LysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLysGluIleSerLeu 228
          :::               ||||||   ::::::   :::      :::
Db   2118 CGTGGACCACATGACCAAGACAGGAAGCTTAGTACCAAAGAGGCCTTGGATGAAGATGTA 2177

Qy    229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
          |||   :::|||      :::|||||||||||||||      ||||||||:::   |||
Db   2178 CCTCCAATTTCCTTTTGGCGTATCCTAAAGTTGAATTCATCTGAATGGCCATATTTTGTG 2237

Qy    249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
          :::|||      :::::::::||||||   :::      |||   |||||||||||||||
Db   2238 GTTGGTATATTCTGTGCCATAGTAAATGGAGCCTTGCAACCAGCATTCTCAATAATATTT 2297

Qy    269 AlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAla 287
          :::|||::::::   :::|||      |||   |||      |||   :::||||||:::
Db   2298 TCTAAGGTTGTAGGGGTTTTCACAAGAAATACTGATGATGAAACCAAACGACATGATAGC 2357

Qy    288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
             ::::::|||::::::|||:::||||||||||||||   |||:::::::::|||:::
Db   2358 AACTTGTTTTCCTTATTGTTTCTGATCCTTGGGGTCATTTCTTTTATTACATTTTTTCTT 2417

Qy    308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
          ||||||   :::|||:::||||||||||||:::   ||||||||||||:::::::
Db   2418 CAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGACTCCGATACATGGTT 2477

Qy    328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
          ||||||:::||||||   ||||||::::::|||||||||      :::|||:::||||||
Db   2478 TTCAAATCCATGCTGAGACAGGATGTGAGCTGGTTTGATAACCCTAAAAACACCACTGGA 2537

Qy    348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
             ||||||||   ||||||   |||      |||:::::::|||||||||||||:::|||
Db   2538 GCATTGACCACCAGGCTCGCCAACGATGCTGGTCAAGTTAAAGGGGCTACAGGAGCCAGA 2597

Qy    368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
          :::   |||:::|||||||||      |||:::|||      :::|||||||||   |||
Db   2598 CTTGCTGTCATTACCCAGAATATAGCAAATCTTGGGACAGGAATCATCATATCCCTAATC 2657

Qy    388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
          ||||||||||:::::::|||   |||:::|||:::|||   |||::::::|||:::   |||
Db   2658 TATGGCTGGCAGTTGACACTTCTACTCTTAGCAATTGTCCCCATCATTGCAATAGCAGGA 2717

Qy    408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
          ::::::|||         ::::::|||   |||   |||||||||:::|||||||:::
Db   2718 GTGGTTGAGATGAAAATGTTGTCTGGACAAGCACTAAAAGATAAGAAGGAGCTAGAAGGT 2777
```

```
Qy     428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
           :::||||||||||||||||||||||:::||||||   ||||||:::||||||||||||||
Db    2778 TCTGGGAAGATTGCTACTGAAGCAATAGAGAACTTCCGCACTGTCGTCTCTTTGACTCGG 2837

Qy     448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
           |||:::   ||||||   ||||||   :::   ||||||      :::||||||
Db    2838 GAGCAGAAGTTTGAAAATATGTATGCCCAGAGCTTGCAGATACCATACAGAAATGCTCTG 2897

Qy     468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
           |||||||||   :::   |||      ::::::|||:::   |||   :::||||||:::
Db    2898 AAGAAAGCACACGTCTTTGGGATCACATTCTCCTTCACCCAGGCCATGATGTATTTTTCG 2957

Qy     488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
           |||||||||   |||||||||||||||||||||:::               ||||||   |||
Db    2958 TATGCTGCTTGTTTCCGGTTTGGTGCTTACTTGGTGGCACGGGAATTAATGACATTTGAA 3017

Qy     508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
            :::   :::||||||:::||||||   :::||||||||||||:::|||:::
Db    3018 AATGTTCTATTAGTATTCTCAGCTATTGTCTTTGGTGCCATGGCAGTTGGGCAGGTCAGT 3077

Qy     528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
                    ||||||:::|||:::||||||||      |||:::|||:::      ::::::
Db    3078 TCATTTGCTCCTGACTATGCCAAAGCCAAAGTGTCGGCATCGCATATCATCATGATCATT 3137

Qy     548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
           ||||||   |||:::|||||||||   |||         |||   ||||||:::|||   |||
Db    3138 GAAAAAGTCCCTTCCATTGACAGCTACAGCACCGGAGGCCTGAAGCCTAATACATTGGAA 3197

Qy     568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
           ||||||::::::|||   ||||||   |||   ||||||   |||||||||:::   :::
Db    3198 GGAAATGTGAAATTTAATGAAGTCGTGTTCAACTACCCCACCCGGCCTGACATCCCAGTG 3257

Qy     588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
           |||:::||||||:::|||   :::::::::|||:::|||:::|||   ||||||||||||
Db    3258 CTTCAGGGGCTGAACCTGGAGGTGAAGAAGGGCCAGACGCTGGCCCTGGTGGGCAGCAGT 3317

Qy     608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
           ||||||||||||||||||   |||||||||:::|||   |||||||||:::   |||
Db    3318 GGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAGCGGTTCTACGACCCCATGGCCGGA 3377

Qy     628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
              |||      ||||||   :::         :::|||||||||||||||||||||:::
Db    3378 ACAGTGTTTCTAGATGGCAAAGAAGTAAACCAACTAAATGTCCAGTGGCTCCGAGCACAC 3437

Qy     648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
           :::   ||||||   ||||||||||:::||||||:::||||||||||||||||||||||||
Db    3438 CTAGGCATTGTGTCCCAAGAGCCCATCCTGTTTGACTGCAGCATCGCTGAGAACATCGCT 3497

Qy     668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
           |||||||||||||||||||||||||||      |||||||||:::   ||||||   |||
Db    3498 TATGGAGACAACAGCCGGGTCGTGTCCCAGGATGAGATCGAGAGGGCGGCGAAGGAAGCC 3557

Qy     688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
           |||||||||   |||||||||   |||||||:::|||||||||||||:::||||||   |||
Db    3558 AACATCCACCAGTTCATTGAGTCACTGCCTGACAAATACAACACCAGAGTAGGAGACAAA 3617

Qy     708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
           |||   ||||||||||||||||||||||||||||:::||||||||||||||||||::::::
Db    3618 GGGACTCAGCTGTCGGGTGGGCAGAAACAACGCATCGCCATAGCACGCGCCCTCGTCAGA 3677

Qy     728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
           :::|||   ||||||||||||||||||||||||||||||||||||   :::||||||||||
Db    3678 CAGCCTCACATTTTGCTCCTGGATGAAGCAACATCAGCCCTGGATACAGAGAGTGAAAAG 3737

Qy     748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
           |||||||||   |||||||||||||||||||   ||||||||||||:::|||:::   |||
Db    3738 GTTGTCCAGGAAGCTCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATCGCTCAC 3797

Qy     768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
           |||||||||   |||||||||||||||||||||:::      |||||||||:::||||||
Db    3798 CGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATTCAGAATGGCAAGGTCAAG 3857

Qy     788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
           |||   |||||||||||:::||||||   :::   |||||||||   :::|||:::
Db    3858 GAGCATGGCACCCACCAGCAGCTGCTGGCACAGAAAGGCATCTATTTCTCCATGGTCAGT 3917

Qy     808 AlaGlnSer 810
              |||:::
Db    3918 GTGCAGGCT 3926
```

```
RESULT 11
OAU78609
LOCUS       OAU78609                3858 bp    mRNA    linear   MAM 11-OCT-2001
DEFINITION  Ovis aries multidrug resistance protein-1 (mdr1) mRNA, complete
            cds.
ACCESSION   U78609
VERSION     U78609.1  GI:2149086
KEYWORDS    .
SOURCE      Ovis aries.
  ORGANISM  Ovis aries
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Caprinae; Ovis.
REFERENCE   1  (bases 1 to 3858)
  AUTHORS   Longley,M., Phua,S.H., van Stijn,T.C. and Crawford,A.M.
  TITLE     Isolation and mapping of the first ruminant multidrug resistance
            genes
  JOURNAL   Anim. Genet. 30 (3), 207-210 (1999)
  MEDLINE   99371931
   PUBMED   10442984
REFERENCE   2  (bases 1 to 3858)
  AUTHORS   Longley,M. and Crawford,A.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-NOV-1996) Biochemistry Department, University of
            Otago, Dunedin, New Zealand
FEATURES             Location/Qualifiers
     source          1..3858
                     /organism="Ovis aries"
                     /db_xref="taxon:9940"
     gene            1..3858
                     /gene="mdr1"
     CDS             1..3858
                     /gene="mdr1"
                     /codon_start=1
                     /product="multidrug resistance protein-1"
                     /protein_id="AAB58489.1"
                     /db_xref="GI:2149087"
                     /translation="MDLEGDRNGRAAGGNFLKRDKKRFFSKKDEKKEKRPTVSTFTMF
                     RYSNWLDRLCMVLGTLAAIIHGAGLPLMMLVFGDMTDSFAGAGNLGNITLSNISNTST
                     IDRTEYGKKLEKEMTTYAYYYSGIGAGVLIAAYIQVSFWCLAAGRQVHRIRKQFFHAI
                     MQQEIGWFDVHDVGKLNTRLTNDVSKINEGIGDKIGMFFQAMATFFTGFIIGFTTGWN
                     LTLVILAIIPVLGLSAAIWAKILSSFTDKKLLAYAKAGAVAEEVLAAIKTVIAFGGQK
                     KELERYNKNLEEAKRIGIKKAITANISMGAAFLLIYASYALAFWYGTSLVLSREYSIG
                     QVLTVFFSVLIGAFSIGQASPNIEAFANARGAAYEVFKIIDNKPSIDSYSNTGHKPDN
                     IKGNLEFRNVHFHYPSRNEVKILKGLNLKVGSGQTVALVGNSGCGKSTTVQLMQRLYD
                     PTEGMVSIDGQDIRTINVRYLREIIGVVSQEPVLFATTIAENIRYGREDVTMDEIQKA
                     VKEANAYDFIMKLPNKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSAL
                     GTESEAVVQAALDKARKGRTTIVIAHRLSTVRNADIIAGLDDGVIVEEGSHDELMGKR
                     GIYFKLVTMQTKGNELELENTPGESLSKTDDLYMSSQDSRSSLIRRKSTRRSIRGSQS
                     QDRKLSTEETLDESVPPVSFWRILKLNITEWPYFVVGVFCAIINGALQPAFSVIFSRI
                     IGIFTRNDNDETKRQNSNLFSLLFLILGIISFITFFLQGFTFGKAGEILTRRLRYLVF
                     RSMLRQDVSWFDDPKNTTGALTTRLANDAAQVKGAVGSRLAVITQNIANLGTGIIISL
                     IYGWQLTLLLLAIVPIIAVAGVIEMKMLSGQALKDKKELEGAGKIATEAIENFRTVVS
                     LTREERFEYMYAQSLQVPYRNSLRRAHVFGITFSITQAMMYFSYAGCFRFGAYLVAQG
                     IMEFQDVLLVFSAVVFGAMAVGQVSTFAPDYAKAKVSAAHVINIIEKIPLIDSYSTEG
                     LKPSTVEGSVAFNDVVFNYPTRPDVPVLRGLSLEVKKGQTLALVGSSGCGKSTVVQLL
                     ERFYDPLAGTVFIDGKEVKQLNVQWLRAHMGIVSQEPILFDCSIGENIAYGDNSRVVS
                     QEEIEHAAKEANIHSFIEMLPDKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLL
                     DEATSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGRIKEHGTH
                     QQLLAQKGIYFTMVSVQAGTKRQ"
BASE COUNT     1072 a    853 c   1002 g    931 t
ORIGIN

Alignment Scores:
Pred. No.:              1.74e-175       Length:         3858
Score:                  2468.00         Matches:        475
```

```
Percent Similarity:       76.73%        Conservative:  158
Best Local Similarity:    57.58%        Mismatches:    174
Query Match:              60.51%        Indels:        18
DB:                       4             Gaps:          4

US-09-873-409-2 (1-812) x OAU78609 (1-3858)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db    1372 ATCGACGGACAGGATATCAGGACCATCAACGTAAGGTATCTGCGGGAGATTATCGGGGTG 1431

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db    1432 GTGAGTCAGGAGCCTGTGCTGTTTGCCACCACGATAGCTGAGAACATTCGCTACGGCCGT 1491

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db    1492 GAAGATGTCACCATGGATGAGATTCAGAAAGCTGTGAAGGAAGCCAACGCCTATGACTTT 1551

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db    1552 ATCATGAAACTGCCTAATAAGTTTGACACCCTGGTTGGAGAGAGAGGAGCCCAGCTAAGC 1611

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db    1612 GGTGGACAGAAGCAAAGAATCGCCATTGCCCGGGCCCTGGTTCGCAACCCCAAGATCCTC 1671

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db    1672 CTGTTGGATGAGGCCACCTCAGCGCTGGGCACCGAGAGTGAAGCGGTGGTTCAGGCGGCC 1731

Qy     122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db    1732 CTGGATAAGGCCAGAAAAGGCCGGACCACCATTGTGATAGCGCATCGTCTGTCCACGGTT 1791

Qy     142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db    1792 CGCAATGCTGACATCATTGCTGGTCTTGATGATGGAGTCATCGTGGAGGAAGGCAGTCAT 1851

Qy     162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLys 181
Db    1852 GATGAGCTCATGGGGAAAAGAGGCATTTACTTCAAACTTGTCACAATGCAG---ACAAAA 1908

Qy     182 LysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu 201
Db    1909 GGAAATGAACTTGAATTGGAAAATACCCCTGGTGAATCCCTAAGTAAAACTGATGACTTG 1968

Qy     202 ProLeuHisSerValLysSerIleLysSerAspPheIle--------------------- 214
Db    1969 TACATGTCCTCCCAAGATTCA---AGATCCAGTCTAATCAGAAGAAAATCAACTCGCAGG 2025

Qy     215 ------------------------AspLysAlaGluGluSerThrGlnSerLysGluIle 226
Db    2026 AGTATCCGTGGATCCCAAAGCCAGGACAGAAAGCTCAGTACAGAAGAAACTTTGGATGAA 2085

Qy     227 SerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPhe 246
Db    2086 AGTGTACCTCCCGTTTCCTTTTGGAGGATTCTGAAGCTGAATATAACTGAATGGCCTTAT 2145

Qy     247 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 266
Db    2146 TTTGTGGTTGGTGTATTTGCGCCATTATAAATGGAGCTCTGCAACCAGCATTCTCAGTC 2205

Qy     267 IlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys---ThrThrLeuLysHis 285
Db    2206 ATATTTTCAAGGATTATAGGAATCTTCACCAGAAATGACAATGACGAAACCAAACGACAG 2265

Qy     286 AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr 305
Db    2266 AACAGTAACTTGTTTTCACTGTTGTTTCTAATCCTTGGAATTATTTCTTTTATTACATTT 2325

Qy     306 PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHis 325
Db    2326 TTCCTTCAGGGCTTCACGTTCGGCAAAGCAGGCGAGATCCTCACCAGGCGGCTGCGGTAC 2385

Qy     326 LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSer 345
Db    2386 CTGGTTTTCAGGTCGATGCTGAGACAGGATGTGAGCTGGTTCGATGACCCTAAAAACACC 2445

Qy     346 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365
Db    2446 ACGGGAGCATTGACGACCAGGCTTGCCAATGACGCCGCTCAAGTTAAAGGGGCTGTAGGT 2505

Qy     366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385
Db    2506 TCCAGGCTTGCTGTAATTACCCAGAATATTGCAAATCTTGGCACAGGCATTATCATATCC 2565

Qy     386 PheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405
Db    2566 CTCATCTATGGCTGGCAGTTAACACTTCTCCTGTTAGCAATTGTACCCATCATTGCAGTA 2625

Qy     406 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425
Db    2626 GCAGGAGTTATTGAAATGAAAATGCTGTCTGGACAAGCCCTGAAAGATAAGAAAGAACTG 2685

Qy     426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445
Db    2686 GAAGGTGCTGGGAAGATTGCCACTGAAGCAATAGAAAACTTCCGGACGGTTGTTTCTTTG 2745

Qy     446 ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn 465
Db    2746 ACTCGGGAGGAGAGGTTTGAATATATGTATGCCCAGAGTTTGCAGGTACCATACAGAAAC 2805

Qy     466 ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr 485
Db    2806 TCCTTGAGGAGAGCGCATGTCTTTGGAATTACATTTTCCATCACTCAGGCAATGATGTAT 2865

Qy     486 PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr 505
Db    2866 TTCTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTGGCCCAAGGCATCATGGAG 2925

Qy     506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys 525
Db    2926 TTTCAGGATGTTCTCTTAGTGTTCTCAGCTGTTGTCTTTGGTGCCATGGCTGTGGGACAG 2985

Qy     526 ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAla 545
Db    2986 GTCAGTACATTTGCTCCTGACTATGCCAAGGCCAAAGTGTCAGCAGCCCACGTCATCAAT 3045

Qy     546 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 565
Db    3046 ATCATTGAAAAAATTCCTCTGATTGACAGCTACAGCACGGAAGGCCTAAAGCCGAGTACA 3105

Qy     566 CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspVal 585
Db    3106 GTGGAAGGAAGTGTGGCATTTAATGACGTCGTGTTCAACTACCCCACTCGACCAGACGTC 3165

Qy     586 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605
Db    3166 CCAGTGCTTCGGGGGCTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCGCTGGTGGGC 3225

Qy     606 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProVal 625
Db    3226 AGTAGTGGCTGTGGGAAGAGCACGGTCGTCCAGCTCCTGGAGCGGTTCTATGACCCCTTG 3285

Qy     626 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArg 645
Db    3286 GCTGGCACAGTGTTTATTGACGGCAAGGAAGTAAAGCAGCTGAATGTCCAGTGGCTGCGG 3345

Qy     646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665
Db    3346 GCACACATGGGCATCGTGTCCCAGGAGCCCATTCTGTTTGACTGCAGCATCGGCGAGAAT 3405

Qy     666 IleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn 685
Db    3406 ATCGCCTACGGAGACAACAGCCGGGTCGTGTCCCAGGAGGAGATTGAGCACGCGGCCAAG 3465
```

```
Qy     686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 705
              |||||||||||||||||||||||   ||||||:::|||||||||||:::||||||
Db    3466 GAGGCCAACATCCACTCCTTCATCGAGATGCTGCCGGACAAATACAACACCAGAGTAGGG 3525

Qy     706 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725
              |||||  ||||||||||||||||||||||||||:::||||||||||||||||
Db    3526 GACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAGCGCATTGCCATAGCTCGCGCCCTT 3585

Qy     726 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745
           :::::::::|||  ||||||||||||||||||||||||||||||||||    :::|||
Db    3586 GTTAGACAGCCTCACATCTTGCTTCTGGATGAAGCTACCTCGGCTCTGGACACGGAAAGT 3645

Qy     746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765
           ||||||||||||||   ||||||||||||||||   ||||||||||||:::|||:::
Db    3646 GAAAAGGTGGTCCAAGAAGCCCTGGACAAAGCCCGAGAGGGACGCACCTGCATCGTGATC 3705

Qy     766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785
              |||||||||||  ||||||||||||||||||||||||||       ||||||:::
Db    3706 GCTCACCGGCTGTCCACCATCCAGAACGCAGACTTGATAGTGGTGTTTCAGAATGGCAGA 3765

Qy     786 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 805
           ||||||||  ||||||||||:::||||||       :::   |||||||||    :::
Db    3766 ATCAAGGAGCATGGCACACACCAGCAGCTGCTGGCGCAGAAAGGCATCTATTTCACCATG 3825

Qy     806 ValAsnAlaGlnSer 810
           |||:::   |||:::
Db    3826 GTCAGTGTCCAGGCT 3840

RESULT 12
MUSMDRA
LOCUS       MUSMDRA                 4084 bp    mRNA    linear   ROD 27-APR-1993
DEFINITION  Mouse mdr gene encoding a multidrug resistance protein mRNA,
            complete cds.
ACCESSION   J03398
VERSION     J03398.1  GI:199109
KEYWORDS    multidrug resistance protein.
SOURCE      Mouse (strain BALB/c) pre-B-cell (cell line 702/3), cDNA to mRNA,
            clone MDR2.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4084)
  AUTHORS   Gros,P., Raymond,M., Bell,J. and Housman,D.
  TITLE     Cloning and characterization of a second member of the mouse mdr
            gene family
  JOURNAL   Mol. Cell. Biol. 8 (7), 2770-2778 (1988)
  MEDLINE   88302195
   PUBMED   3405218
REFERENCE   2  (sites)
  AUTHORS   Raymond,M. and Gros,P.
  TITLE     Mammalian multidrug-resistance gene: correlation of exon
            organization with structural domains and duplication of an
            ancestral gene
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 86 (17), 6488-6492 (1989)
  MEDLINE   89367274
   PUBMED   2570420
REFERENCE   3  (sites)
  AUTHORS   Raymond,M. and Gros,P.
  TITLE     Cell-specific activity of cis-acting regulatory elements in the
            promoter of the mouse multidrug resistance gene mdr1
  JOURNAL   Mol. Cell. Biol. 10 (11), 6036-6040 (1990)
  MEDLINE   91042535
   PUBMED   2248681
COMMENT     Submitted in computer readable form by P.Gros 08-JUL-1988.
FEATURES             Location/Qualifiers
     source          1..4084
                     /organism="Mus musculus"
                     /strain="BALB/c"
                     /sub_species="domesticus"
                     /db_xref="taxon:10090"
                     /cell_line="702/3"
                     /cell_type="pre-B cell"
     CDS             194..4024
                     /codon_start=1
                     /product="multidrug resistance protein"
                     /protein_id="AAA39516.1"
                     /db_xref="GI:387428"
                     /translation="MDLEAARNGTARRLDGDFELGSISNQGREKKKKVNLIGLLTLFR
                     YSDWQDKLFMFLGTLMAIAHGSGLPLMMIVFGEMTDKFVDNTGNFSLPVNFSLSMLNP
                     GRILEEEMTRYAYYYSGLGGGVLVAAYIQVSFWTLAAGRQIKKIRQKFFHAILRQEMG
                     WFDIKGTTELNTRLTDDVSKISEGIGDKVGMFFQAIATFFAGFIVGFIRGWKLTLVIM
                     AISPILGLSTAVWAKILSTFSDKELAAYAKAGAVAEEAPGAIRTVIAFGGQNKELERY
                     QKHLENAKKIGIKKAISANISMGIAFLLIYASYALAFWYGSTLVISKEYTIGNAMTVF
                     FSILIGAFSVGQAAPCIDAFANARGAAYVIFDIIDNNPKIDSFSERGHKPDNIKGNLE
                     FSDVHFSYPSRANIKILKGLNLKVKSGQTVALVGNSGCGKSTTVQLLQRLYDPTEGKI
                     SIDGQDIRNFNVRCLREIIGVVSQEPVLFSTTIAENIRYGRGNVTMDEIEKAVKEANA
                     YDFIMKLPQKFDTLVGDRGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTESEA
                     EVQAALDKAREGRTTIVIAHRLSTIRNADVIAGFEDGVIVEQGSHSELMKKEGIYFRL
                     VNMQTAGSQILSEEFEVELSDEKAAGDVAPNGWKARIFRNSTKKSLKSPHQNRLDEET
                     NELDANVPPVSFLKVLKLNKTEWPYFVVGTVCAIANGALQPAFSIILSEMIAIFGPGD
                     DAVKQQKCNMFSLVFLGLGVLSFFTFFLQGFTFGKAGEILTTRLRSMAFKAMLRQDMS
                     WFDDHKNSTGALSTRLATDAAQVQGATGTKLALIAQNTANLGTGIIISFIYGWQLTLL
                     LLSVVPFIAVAGIVEMKMLAGNAKRDKKEMEAAGKIATEAIENIRTVVSLTQERKFES
                     MYVEKLHGPYRNSVRKAHIYGITFSISQAFMYFSYAGCFRFGSYLIVNGHMRFKDVIL
                     VFSAIVLGAVALGHASSFAPDYAKAKLSAAYLFSLFERQPLIDSYSGEGLWPDKFEGS
                     VTFNEVVFNYPTRANVPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERFYDPMAG
                     SVLLDGQEAKKLNVQWLRAQLGIVSQEPILFDCSIAENIAYGDNSRVVPHDEIVRAAK
                     EANIHPFIETLPQKYNTRVGDKGTQLSGGQKQRIAIARALIRQPRVLLLDEATSALDT
                     ESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVIENGKVKEHGTHQQLLAQKGI
                     YFSMVNIQAGTQNL"
BASE COUNT     1117 a    934 c   1075 g    958 t
ORIGIN      43 bp upstream of HinfI site.

Alignment Scores:
Pred. No.:              2.05e-175      Length:          4084
Score:                  2467.50        Matches:         475
Percent Similarity:     76.46%         Conservative:    152
Best Local Similarity:  57.93%         Mismatches:      182
Query Match:            60.49%         Indels:          11
DB:                     10             Gaps:            1

US-09-873-409-2 (1-812) x MUSMDRA (1-4084)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
           :::|||      ||||||||      |||||||||      |||:::   |||||||||
Db    1547 ATCGATGGGCAGGATATCAGGAACTTTAACGTCAGGTGTCTAAGGGAAATCATTGGTGTG 1606

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
              ||||||||||||||||||||||||      ||||||:::      ||||||:::|||||||||
Db    1607 GTAAGTCAAGAGCCCGTGCTGTTCTCTACTACGATCGCTGAAAATATCCGCTATGGCCGT 1666

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
              :::||||||   :::|||:::|||:::|||   :::||||||||||||||||||||||
Db    1667 GGGAATGTAACGATGGATGAGATTGAGAAAGCCGTCAAAGAGGCCAATGCCTATGACTTC 1726

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           ||||||:::    |||    ||||||:::||||||||||||::::::|||||||||:::|||
Db    1727 ATCATGAAACTGCCCCAGAAATTTGACACCCTGGTTGGTGATAGAGGGGCGCAGCTGAGT 1786

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1787 GGGGGACAGAAACAGAGAATCGCCATTGCCCGGGCCCTAGTACGCAACCCCAAGATCCTC 1846

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
           :::||||||||||||||||||||||||||||||:::||||||::::::   ||||||||||||
Db    1847 CTGCTGGACGAGGCCACCTCAGCCCTGGACACTGAAAGTGAAGCTGAGGTGCAGGCCGCA 1906

Qy     122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
           |||:::||||||   :::||||||||||||||||||:::||||||||||||||||||
Db    1907 CTGGATAAGGCCAGAGAAGGCCGAACCACCATTGTGATAGCTCACCGATTGTCTACCATC 1966

Qy     142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
           |||:::||||||:::|||         :::||||||::::::   |||:::|||:::|||
Db    1967 CGGAACGCAGATGTCATCGCTGGGTTTGAGGATGGAGTCATTGTGGAACAAGGAAGTCAC 2026
```

```
Qy    162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178
Db   2027 AGTGAGCTGATGAAGAAGGAAGGGATCTACTTCAGACTCGTTAACATGCAGACAGCAGGA 2086

Qy    179 -----------------------AspIleLysLysAlaAspGluGlnMetGluSerMet 190
Db   2087 AGCCAGATCCTGTCAGAAGAATTTGAAGTTGAGCTAAGTGACGAAAAGGCTGCTGGAGAT 2146

Qy    191 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 210
Db   2147 GTGGCCCCAAATGGCTGGAAAGCACGCATATTTAGGAATTCTACAAAGAAAAGTCTTAAA 2206

Qy    211 SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGlu 230
Db   2207 AGTCCACATCAGAATAGGCTGGATGAAGAAACCAATGAACTTGATGCAAACGTGCCACCA 2266

Qy    231 ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGly 250
Db   2267 GTGTCTTTTCTGAAGGTCTTAAAACTGAATAAAACAGAGTGGCCCTACTTTGTGGTGGGA 2326

Qy    251 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 270
Db   2327 ACAGTCTGTGCCATTGCCAATGGAGCCCTCCAGCCGGCTTTCTCCATCATCCTGTCTGAG 2386

Qy    271 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 290
Db   2387 ATGATAGCTATCTTTGGCCCTGGGGATGACGCAGTGAAGCAGCAAAAGTGTAACATGTTC 2446

Qy    291 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 310
Db   2447 TCCCTGGTCTTCTTGGGCCTAGGAGTCCTCTCCTTCTTTACTTTCTTCCTTCAGGGCTTC 2506

Qy    311 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 330
Db   2507 ACGTTTGGGAAAGCTGGAGAGATCCTCACCACAAGGCTCCGGTCCATGGCCTTTAAAGCG 2566

Qy    331 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 350
Db   2567 ATGCTAAGGCAGGACATGAGCTGGTTTGATGATCATAAAAACAGTACTGGAGCACTTTCT 2626

Qy    351 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 370
Db   2627 ACAAGACTCGCCACAGATGCTGCGCAAGTCCAAGGAGCCACGGGAACCAAGTTGGCTTTA 2686

Qy    371 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrp 390
Db   2687 ATTGCACAGAACACAGCCAACCTTGGAACCGGTATTATTATATCATTTATTTACGGTTGG 2746

Qy    391 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 410
Db   2747 CAACTGACACTTCTGCTGTTATCGGTTGTTCCATTCATTGCTGTAGCAGGAATTGTTGAA 2806

Qy    411 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 430
Db   2807 ATGAAAATGTTGGCTGGCAATGCCAAGAGAGATAAAAAGGAAATGGAAGCTGCTGGAAAG 2866

Qy    431 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 450
Db   2867 ATTGCAACAGAGGCAATAGAAAATATTCGAACTGTTGTATCCTTGACCCAAGAAAGAAAA 2926

Qy    451 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 470
Db   2927 TTTGAGTCAATGTATGTTGAAAAATTGCATGGACCTTACAGGAATTCGGTGCGGAAGGCA 2986

Qy    471 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 490
Db   2987 CACATCTACGGCATCACTTTTAGCATCTCCCAAGCATTCATGTATTTTTCTTATGCTGGC 3046

Qy    491 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 510
Db   3047 TGTTTTCGATTTGGTTCTTACCTAATTGTGAATGGACATATGCGCTTCAAAGATGTCATT 3106

Qy    511 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 530
Db   3107 CTGGTCTTTTCTGCAATTGTGCTTGGCGCGGTGGCTCTAGGACACGCCAGCTCATTTGCT 3166

Qy    531 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 550
Db   3167 CCGGACTATGCAAAAGCCAAGCTGTCTGCAGCATACTTGTTCAGCCTGTTTGAAAGACAA 3226

Qy    551 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 570
Db   3227 CCTCTGATTGACAGCTACAGTGGAGAAGGGCTGTGGCCTGATAAGTTTGAAGGAAGCGTG 3286

Qy    571 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 590
Db   3287 ACATTTAATGAAGTCGTGTTCAACTATCCCACCCGGGCCAACGTGCCAGTGCTTCAGGGG 3346

Qy    591 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 610
Db   3347 CTGAGCCTTGAGGTGAAGAAGGGCCAGACGCTGGCCCTGGTGGGCAGCAGTGGCTGCGGG 3406

Qy    611 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 630
Db   3407 AAGAGCACAGTGGTCCAGCTGCTCGAGCGCTTCTATGACCCCATGGCTGGATCAGTGCTC 3466

Qy    631 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 650
Db   3467 TTAGATGGTCAAGAAGCAAAGAAACTCAATGTCCAGTGGCTCCGAGCTCAACTGGGCATT 3526

Qy    651 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 670
Db   3527 GTGTCCCAGGAACCCATTCTCTTTGACTGCAGCATCGCAGAGAACATCGCCTATGGAGAC 3586

Qy    671 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 690
Db   3587 AACAGCCGGGTCGTGCCTCATGATGAGATTGTGAGGGCAGCCAAGGAGGCCAACATCCAC 3646

Qy    691 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 710
Db   3647 CCCTTCATCGAGACGCTGCCCCAAAAATATAACACAAGAGTAGGAGACAAGGGGACGCAG 3706

Qy    711 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 730
Db   3707 CTCTCTGGGGGCCAGAAGCAGAGGATTGCCATCGCCCGAGCCCTCATCAGACAGCCTCGG 3766

Qy    731 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 750
Db   3767 GTCCTACTGCTGGATGAAGCCACGTCAGCTCTGGATACTGAGAGTGAAAAGGTTGTCCAG 3826

Qy    751 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 770
Db   3827 GAAGCACTGGACAAAGCCAGGGAAGGCCGCACCTGCATTGTGATCGCTCACCGCCTGTCC 3886

Qy    771 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 790
Db   3887 ACCATCCAGAACGCGGACTTGATCGTGGTGATTGAGAACGGCAAGGTCAAGGAGCACGGC 3946

Qy    791 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 810
Db   3947 ACCCACCAGCAGCTGCTGGCGCAGAAGGGCATCTATTTCTCAATGGTCAACATCCAGGCC 4006
```

```
RESULT 13
AX024455
LOCUS       AX024455                3924 bp    DNA     linear   PAT 15-SEP-2000
DEFINITION  Sequence 2 from Patent DE19857750.
ACCESSION   AX024455
VERSION     AX024455.1  GI:10184623
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3924)
  AUTHORS   Stremmel,W.
  JOURNAL   Patent: DE 19857750-A 2 24-FEB-2000;
```

```
              STREMMEL WOLFGANG (DE)
FEATURES                 Location/Qualifiers
     source              1..3924
                         /organism="Homo sapiens"
                         /db_xref="taxon:9606"
                         /note="MDR3/menschliche MDR-Genfamilie"
BASE COUNT     1145 a    790 c    977 g   1012 t
ORIGIN

Alignment Scores:
Pred. No.:               2.74e-175      Length:         3924
Score:                   2465.50        Matches:        478
Percent Similarity:      76.46%         Conservative:   149
Best Local Similarity:   58.29%         Mismatches:     182
Query Match:             60.44%         Indels:         11
DB:                      6              Gaps:           3

US-09-873-409-2 (1-812) x AX024455 (1-3924)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db    1395 ATTGATGGGCAGGATATTAGGAACTTTAATGTAAACTATCTGAGGGAAATCATTGGTGTG 1454

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db    1455 GTGAGTCAGGAGCCGGTGCTGTTTTCCACCACAATTGCTGAAAATATTTGTTATGGCCGT 1514

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db    1515 GGAAATGTAACCATGGATGAGATAAAGAAAGCTGTCAAAGAGGCCAACGCCTATGAGTTT 1574

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db    1575 ATCATGAAATTACCACAGAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGCTGAGT 1634

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db    1635 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTT 1694

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db    1695 CTGCTGGATGAGGCCACGTCAGCATTGGACACAGAAAGTGAAGCTGAGGTACAGGCAGCT 1754

Qy     122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db    1755 CTGGATAAGGCCAGAGAAGGCCGGACCACCATTGTGATAGCACACCGACTGTCTACGGTC 1814

Qy     142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db    1815 CGAAATGCAGATGTCATCGCTGGGTTTGAGGATGGAGTAATTGTGGAGCAAGGAAGCCAC 1874

Qy     162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal------------------ 175
Db    1875 AGCGAACTGATGAAGAAGGAAGGGGTGTACTTCAAACTTGTCAACATGCAGACATCAGGA 1934

Qy     176 ---------MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyr 192
Db    1935 AGCCAGATCCAGTCAGAAGAATTTGAACTAAATGATGAAAAGGCTGCCACTAGAATGGCC 1994

Qy     193 SerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys---Ser 211
Db    1995 CCAAATGGCTGGAAATCTCGCCTATTTAGGCATTCTACTCAGAAAAACCTTAAAAATTCA 2054

Qy     212 AspPheIleAspLysAla---GluGluSerThrGlnSerLysGluIleSerLeuProGlu 230
Db    2055 CAAATGTGTCAGAAGAGCCTTGATGTGGAAACCGATGGACTTGAAGCAAATGTGCCACCA 2114

Qy     231 ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGly 250
Db    2115 GTGTCCTTTCTGAAGGTCCTGAAACTGAATAAAACAGAATGGCCCTACTTTGTCGTGGGA 2174

Qy     251 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 270
Db    2175 ACAGTATGTGCCATTGCCAATGGGGGGCTTCAGCCGGCATTTTCAGTCATATTCTCAGAG 2234

Qy     271 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 290
Db    2235 ATCATAGCGATTTTTGGACCAGGCGATGATGCAGTGAAGCAGCAGAAGTGCAACATATTC 2294

Qy     291 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 310
Db    2295 TCTTTGATTTTCTTATTTCTGGGAATTATTTCTTTTTTTACTTTCTTCCTTCAGGGTTTC 2354

Qy     311 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 330
Db    2355 ACGTTTGGGAAAGCTGGCGAGATCCTCACCAGAAGACTGCGGTCAATGGCTTTTAAAGCA 2414

Qy     331 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 350
Db    2415 ATGCTAAGACAGGACATGAGCTGGTTTGATGACCATAAAAACAGTACTGGTGCACTTTCT 2474

Qy     351 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 370
Db    2475 ACAAGACTTGCCACAGATGCTGCCCAAGTCCAAGGAGCCACAGGAACCAGGTTGGCTTTA 2534

Qy     371 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrp 390
Db    2535 ATTGCACAGAATATAGCTAACCTTGGAACTGGTATTATCATATCATTTATCTACGGTTGG 2594

Qy     391 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 410
Db    2595 CAGTTAACCCTATTGCTATTAGCAGTTGTTCCAATTATTGCTGTGTCAGGAATTGTTGAA 2654

Qy     411 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 430
Db    2655 ATGAAATTGTTGGCTGGAAATGCCAAAAGAGATAAAAAAGAACTGGAAGCTGCTGGAAAG 2714

Qy     431 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 450
Db    2715 ATTGCAACAGAGGCAATAGAAAATATTAGGACAGTTGTGTCTTTGACCCAGGAAAGAAAA 2774

Qy     451 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 470
Db    2775 TTTGAATCAATGTATGTTGAAAAATTGTATGGACCTTACAGGAATTCTGTGCAGAAGGCA 2834

Qy     471 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 490
Db    2835 CACATCTATGGAATTACTTTTAGTATCTCACAAGCATTTATGTATTTTTCCTATGCCGGT 2894

Qy     491 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 510
Db    2895 TGTTTTCGATTTGGTGCATATCTCATTGTGAATGGACATATGCGCTTCAGAGATGTTATT 2954

Qy     511 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 530
Db    2955 CTGGTGTTTTCTGCAATTGTATTTGGTGCAGTGGCTCTAGGACATGCCAGTTCATTTGCT 3014

Qy     531 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 550
Db    3015 CCAGACTATGCTAAAGCTAAGCTGTCTGCAGCCCACTTATTCATGCTGTTTGAAAGACAA 3074

Qy     551 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 570
Db    3075 CCTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAGCCTGATAAATTTGAAGGAAATATA 3134

Qy     571 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 590
Db    3135 ACATTTAATGAAGTCGTGTTCAACTATCCCACCCGAGCAAACGTGCCAGTGCTTCAGGGG 3194

Qy     591 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 610
Db    3195 CTGAGCCTGGAGGTGAAGAAAGGCCAGACACTAGCCCTGGTGGGCAGCAGTGGCTGTGGG 3254

Qy     611 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 630
Db    3255 AAGAGCACGGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCGGGGACAGTGCTT 3314
```

```
Qy     631 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 650
              |||||   :::|||||:::||||||||||||||||||:::|||:::    |||
Db    3315 CTCGATGGTCAAGAAGCAAAGAAACTCAATGTCCAGTGGCTCAGAGCTCAACTCGGAATC 3374

Qy     651 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 670
           |||   ||||||||:::|||||:::||||||||||||||||||||||||||||
Db    3375 GTGTCTCAGGAGCCTATCCTATTTGACTGCAGCATTGCCGAGAATATTGCCTATGGAGAC 3434

Qy     671 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 690
           ||||||||||||||        ||||||||      ||||||   ||||||||||||||
Db    3435 AACAGCCGGGTTGTATCACAGGATGAAATTGTGAGTGCAGCCAAAGCTGCCAACATACAT 3494

Qy     691 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 710
              |||||||||      ||||||        ||||||    |||::|||||    ||||||     |||
Db    3495 CCTTTCATCGAGACGTTACCCCACAAATATGAAACAAGAGTGGGAGATAAGGGGACTCAG 3554

Qy     711 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 730
           ||||||||||||||||||||||||:::|||||||||||||||||||:::::::::|||:::
Db    3555 CTCTCAGGAGGTCAAAAACAGAGGATTGCTATTGCCCGAGCCCTCATCAGACAACCTCAA 3614

Qy     731 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 750
           ||||||||||||||||||||||||||||||||||||      :::||||||||||||||||
Db    3615 ATCCTCCTGTTGGATGAAGCTACATCAGCTCTGGATACTGAAAGTGAAAAGGTTGTCCAA 3674

Qy     751 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 770
              ||||||||||||||||     ||||||||||:::|||:::    |||||||||||
Db    3675 GAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCC 3734

Qy     771 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 790
              ||||||||||||||||||||||||||      ||||||:::::|||||    |||
Db    3735 ACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGGAGAGTCAAGGAGCATGGC 3794

Qy     791 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 810
           ||||||||:::||||||        :::    |||||||||    :::|||:::    |||:::
Db    3795 ACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3854
```

```
RESULT 14
AX409652
LOCUS       AX409652               3924 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 2299 from Patent WO0229103.
ACCESSION   AX409652
VERSION     AX409652.1  GI:21442357
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
  TITLE     Gene expression profiles in liver cancer
  JOURNAL   Patent: WO 0229103-A 2299 11-APR-2002;
            GENE LOGIC INC (US)
FEATURES             Location/Qualifiers
     source          1..3924
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="EMBL/GenBank Accession No. M23234"
BASE COUNT     1145 a    790 c    977 g   1012 t
ORIGIN

Alignment Scores:
Pred. No.:              2.74e-175    Length:          3924
Score:                  2465.50      Matches:         478
Percent Similarity:     76.46%       Conservative:    149
Best Local Similarity:  58.29%       Mismatches:      182
Query Match:            60.44%       Indels:          11
DB:                     6            Gaps:            3

US-09-873-409-2 (1-812) x AX409652 (1-3924)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
             :::|||      |||||||||      ||||||    :::    |||:::    |||||||||
Db    1395 ATTGATGGGCAGGATATTAGGAACTTTAATGTAAACTATCTGAGGGAAATCATTGGTGTG 1454

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           ||||||||||||||||||||||||   |||||||||:::    ||||||   ||||||||
Db    1455 GTGAGTCAGGAGCCGGTGCTGTTTTCCACCACAATTGCTGAAAATATTTGTTATGGCCGT 1514

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
              :::||||||    :::|||:::::::::|||    :::||||||||||||||||:::|||
Db    1515 GGAAATGTAACCATGGATGAGATAAAGAAAGCTGTCAAAGAGGCCAACGCCTATGAGTTT 1574

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           ||||||:::    |||    ||||||:::|||||||||||||:::||||||||:::|||
Db    1575 ATCATGAAATTACCACAGAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGCTGAGT 1634

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1635 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTT 1694

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
           :::|||||||||||||||||||||||||||:::||||||::::::    ||||||||||
Db    1695 CTGCTGGATGAGGCCACGTCAGCATTGGACACAGAAAGTGAAGCTGAGGTACAGGCAGCT 1754

Qy     122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
           |||:::||||||    :::|||||||||||||||||:::||||||||||||||||||:::
Db    1755 CTGGATAAGGCCAGAGAAGGCCGGACCACCATTGTGATAGCACACCGACTGTCTACGGTC 1814

Qy     142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
           |||:::||||||:::|||         :::|||||::::::    |||:::|||:::|||
Db    1815 CGAAATGCAGATGTCATCGCTGGGTTTGAGGATGGAGTAATTGTGGAGCAAGGAAGCCAC 1874

Qy     162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal------------------ 175
           :::||||||||    |||    |||:::|||:::    ||||||
Db    1875 AGCGAACTGATGAAGAAGGAAGGGGTGTACTTCAAACTTGTCAACATGCAGACATCAGGA 1934

Qy     176 ---------MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyr 192
                    |||::::::    :::         ||||||:::         :::
Db    1935 AGCCAGATCCAGTCAGAAGAATTTGAACTAAATGATGAAAAGGCTGCCACTAGAATGGCC 1994

Qy     193 SerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys---Ser 211
                    |||:::                     :::    |||::::::|||   |||
Db    1995 CCAAATGGCTGGAAATCTCGCCTATTTAGGCATTCTACTCAGAAAAACCTTAAAAATTCA 2054

Qy     212 AspPheIleAspLysAla---GluGluSerThrGlnSerLysGluIleSerLeuProGlu 230
                    |||:::      :::         |||         |||      ::::::|||
Db    2055 CAAATGTGTCAGAAGAGCCTTGATGTGGAAACCGATGGACTTGAAGCAAATGTGCCACCA 2114

Qy     231 ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGly 250
           ||||||   |||||:::|||||||||||||||   ||||||||:::    |||:::|||
Db    2115 GTGTCCTTTCTGAAGGTCCTGAAACTGAATAAAACAGAATGGCCCTACTTTGTCGTGGGA 2174

Qy     251 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 270
           |||:::    ::::::   ||||||     :::    |||   ||||||:::||||||::::::
Db    2175 ACAGTATGTGCCATTGCCAATGGGGGGCTTCAGCCGGCATTTTCAGTCATATTCTCAGAG 2234

Qy     271 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 290
           ||||||   :::||||||         |||               :::         |||:::
Db    2235 ATCATAGCGATTTTTGGACCAGGCGATGATGCAGTGAAGCAGCAGAAGTGCAACATATTC 2294

Qy     291 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 310
           |||:::|||||||:::    ||||||:::|||   |||    ::::::|||:::||||||
Db    2295 TCTTTGATTTTCTTATTTCTGGGAATTATTTCTTTTTTACTTTCTTCCTTCAGGGTTTC 2354

Qy     311 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 330
              :::|||:::||||||||||||||||||||||||       :::||||||||||||
Db    2355 ACGTTTGGGAAAGCTGGCGAGATCCTCACCAGAAGACTGCGGTCAATGGCTTTTAAAGCA 2414

Qy     331 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 350
           ||||||   ||||||::::::||||||||||:::    :::|||||||||||||    |||:::
Db    2415 ATGCTAAGACAGGACATGAGCTGGTTTGATGACCATAAAAACAGTACTGGTGCACTTTCT 2474

Qy     351 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 370
           |||   ||||||   |||   ||||||:::||||||||||||||:::|||:::    :::
Db    2475 ACAAGACTTGCCACAGATGCTGCCCAAGTCCAAGGAGCCACAGGAACCAGGTTGGCTTTA 2534
```

```
Qy      371 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrp 390
Db     2535 ATTGCACAGAATATAGCTAACCTTGGAACTGGTATTATCATATCATTTATCTACGGTTGG 2594

Qy      391 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 410
Db     2595 CAGTTAACCCTATTGCTATTAGCAGTTGTTCCAATTATTGCTGTGTCAGGAATTGTTGAA 2654

Qy      411 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 430
Db     2655 ATGAAATTGTTGGCTGGAAATGCCAAAAGAGATAAAAAAGAACTGGAAGCTGCTGGAAAG 2714

Qy      431 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 450
Db     2715 ATTGCAACAGAGGCAATAGAAAATATTAGGACAGTTGTGTCTTTGACCCAGGAAAGAAAA 2774

Qy      451 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 470
Db     2775 TTTGAATCAATGTATGTTGAAAAATTGTATGGACCTTACAGGAATTCTGTGCAGAAGGCA 2834

Qy      471 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 490
Db     2835 CACATCTATGGAATTACTTTAGTATCTCACAAGCATTTATGTATTTTTCCTATGCCGGT 2894

Qy      491 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 510
Db     2895 TGTTTTCGATTTGGTGCATATCTCATTGTGAATGGACATATGCGCTTCAGAGATGTTATT 2954

Qy      511 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 530
Db     2955 CTGGTGTTTTCTGCAATTGTATTTGGTGCAGTGGCTCTAGGACATGCCAGTTCATTTGCT 3014

Qy      531 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 550
Db     3015 CCAGACTATGCTAAAGCTAAGCTGTCTGCAGCCCACTTATTCATGCTGTTTGAAAGACAA 3074

Qy      551 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 570
Db     3075 CCTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAGCCTGATAAATTTGAAGGAAATATA 3134

Qy      571 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 590
Db     3135 ACATTTAATGAAGTCGTGTTCAACTATCCCACCCGAGCAAACGTGCCAGTGCTTCAGGGG 3194

Qy      591 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 610
Db     3195 CTGAGCCTGGAGGTGAAGAAAGGCCAGACACTAGCCCTGGTGGGCAGCAGTGGCTGTGGG 3254

Qy      611 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 630
Db     3255 AAGAGCACGGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCGGGGACAGTGCTT 3314

Qy      631 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 650
Db     3315 CTCGATGGTCAAGAAGCAAAGAAACTCAATGTCCAGTGGCTCAGAGCTCAACTCGGAATC 3374

Qy      651 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 670
Db     3375 GTGTCTCAGGAGCCTATCCTATTTGACTGCAGCATTGCCGAGAATATTGCCTATGGAGAC 3434

Qy      671 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 690
Db     3435 AACAGCCGGGTTGTATCACAGGATGAAATTGTGAGTGCAGCCAAAGCTGCCAACATACAT 3494

Qy      691 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 710
Db     3495 CCTTTCATCGAGACGTTACCCCACAAATATGAAACAAGAGTGGGAGATAAGGGGACTCAG 3554

Qy      711 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 730
Db     3555 CTCTCAGGAGGTCAAAAACAGAGGATTGCTATTGCCCGAGCCCTCATCAGACAACCTCAA 3614

Qy      731 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 750
Db     3615 ATCCTCCTGTTGGATGAAGCTACATCAGCTCTGGATACTGAAAGTGAAAAGGTTGTCCAA 3674

Qy      751 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 770
Db     3675 GAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCC 3734

Qy      771 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 790
Db     3735 ACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGGAGAGTCAAGGAGCATGGC 3794

Qy      791 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 810
Db     3795 ACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3854
```

```
RESULT 15
HUMMDR3
LOCUS       HUMMDR3                 3924 bp    mRNA    linear   PRI 11-JUN-1993
DEFINITION  Human membrane glycoprotein P (mdr3) mRNA, complete cds.
ACCESSION   M23234
VERSION     M23234.1  GI:187501
KEYWORDS    P-glycoprotein; membrane glycoprotein.
SOURCE      Human liver, cDNA to mRNA, clone 3.27.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3924)
  AUTHORS   van der Bliek,A.M., Kooiman,P.M., Schneider,C. and Borst,P.
  TITLE     Sequence of mdr3 cDNA encoding a human P-glycoprotein
  JOURNAL   Gene 71 (2), 401-411 (1988)
  MEDLINE   89138016
   PUBMED   2906314
COMMENT     Draft entry and computer-readable sequence for [1] kindly provided
            by P. Borst, 21-MAR-1989.
FEATURES             Location/Qualifiers
     source          1..3924
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     mRNA            <1..3924
                     /product="mdr3 mRNA"
     CDS             33..3872
                     /note="P-glycoprotein"
                     /codon_start=1
                     /protein_id="AAA36207.1"
                     /db_xref="GI:307181"
                     /translation="MDLEAAKNGTAWRPTSAEGDFELGISSKQKRKKTKTVKMIGVLT
                     LFRYSDWQDKLFMSLGTIMAIAHGSGLPLMMIVFGEMTDKFVDTAGNFSFPVNFSLSL
                     LNPGKILEEEMTRYAYYYSGLGAGVLVAAYIQVSFWTLAAGRQIRKIRQKFFHAILRQ
                     EIGWFDINDTTELNTRLTDDISKISEGIGDKVGMFFQAVATFFAGFIVGFIRGWKLTL
                     VIMAISPILGLSAAVWAKILSAFSDKELAAYAKAGAVAEEALGAIRTVIAFGGQNKEL
                     ERYQKHLENAKEIGIKKAISANISMGIAFLLIYASYALAFWYGSTLVISKEYTIGNAM
                     TVFFSILIGAFSVGQAAPCIDAFANARGAAYVIFDIIDNNPKIDSFSERGHKPDSIKG
                     NLEFNDVHFSYPSRANVKILKGLNLKVQSGQTVALVGSSGCGKSTTVQLIQRLYDPDE
                     GTINIDGQDIRNFNVNYLREIIGVVSQEPVLFSTTIAENICYGRGNVTMDEIKKAVKE
                     ANAYEFIMKLPQKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTE
                     SEAEVQAALDKAREGRTTIVIAHRLSTVRNADVIAGFEDGVIVEQGSHSELMKKEGVY
                     FKLVNMQTSGSQIQSEEFELNDEKAATRMAPNGWKSRLFRHSTQKNLKNSQMCQKSLD
                     VETDGLEANVPPVSFLKVLKLNKTEWPYFVVGTVCAIANGGLQPAFSVIFSEIIAIFG
                     PGDDAVKQQKCNIFSLIFLFLGIISFFTFFLQGFTFGKAGEILTRRLRSMAFKAMLRQ
                     DMSWFDDHKNSTGALSTRLATDAAQVQGATGTRLALIAQNIANLGTGIIISFIYGWQL
                     TLLLLAVVPIIAVSGIVEMKLLAGNAKRDKKELEAAGKIATEAIENIRTVVSLTQERK
                     FESMYVEKLYGPYRNSVQKAHIYGITFSISQAFMYFSYAGCFRFGAYLIVNGHMRFRD
                     VILVFSAIVFGAVALGHASSFAPDYAKAKLSAAHLFMLFERQPLIDSYSEEGLKPDKF
                     EGNITFNEVVFNYPTRANVPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERFYDP
                     LAGTVLLDGQEAKKLNVQWLRAQLGIVSQEPILFDCSIAENIAYGDNSRVVSQDEIVS
                     AAKAANIHPFIETLPHKYETRVGDKGTQLSGGQKQRIAIARALIRQPQILLLDEATSA
                     LDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGRVKEHGTHQQLLAQ
                     KGIYFSMVSVQAGTQNL"
BASE COUNT     1145 a    790 c    977 g   1012 t
ORIGIN

Alignment Scores:
```

```
Pred. No.:                2.74e-175    Length:          3924
Score:                    2465.50      Matches:         478
Percent Similarity:       76.46%       Conservative:    149
Best Local Similarity:    58.29%       Mismatches:      182
Query Match:              60.44%       Indels:          11
DB:                       9            Gaps:            3

US-09-873-409-2 (1-812) x HUMMDR3 (1-3924)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db   1395 ATTGATGGGCAGGATATTAGGAACTTTAATGTAAACTATCTGAGGGAAATCATTGGTGTG 1454

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db   1455 GTGAGTCAGGAGCCGGTGCTGTTTTCCACCACAATTGCTGAAAATATTTGTTATGGCCGT 1514

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db   1515 GGAAATGTAACCATGGATGAGATAAAGAAAGCTGTCAAAGAGGCCAACGCCTATGAGTTT 1574

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db   1575 ATCATGAAATTACCACAGAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGCTGAGT 1634

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db   1635 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTT 1694

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db   1695 CTGCTGGATGAGGCCACGTCAGCATTGGACACAGAAAGTGAAGCTGAGGTACAGGCAGCT 1754

Qy    122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db   1755 CTGGATAAGGCCAGAGAAGGCCGGACCACCATTGTGATAGCACACCGACTGTCTACGGTC 1814

Qy    142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db   1815 CGAAATGCAGATGTCATCGCTGGGTTTGAGGATGGAGTAATTGTGGAGCAAGGAAGCCAC 1874

Qy    162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal------------------ 175
Db   1875 AGCGAACTGATGAAGAAGGAAGGGGTGTACTTCAAACTTGTCAACATGCAGACATCAGGA 1934

Qy    176 ---------MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyr 192
Db   1935 AGCCAGATCCAGTCAGAAGAATTTGAACTAAATGATGAAAAGGCTGCCACTAGAATGGCC 1994

Qy    193 SerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys---Ser 211
Db   1995 CCAAATGGCTGGAAATCTCGCCTATTTAGGCATTCTACTCAGAAAAACCTTAAAAATTCA 2054

Qy    212 AspPheIleAspLysAla---GluGluSerThrGlnSerLysGluIleSerLeuProGlu 230
Db   2055 CAAATGTGTCAGAAGAGCCTTGATGTGGAAACCGATGGACTTGAAGCAAATGTGCCACCA 2114

Qy    231 ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGly 250
Db   2115 GTGTCCTTTCTGAAGGTCCTGAAACTGAATAAAACAGAATGGCCCTACTTTGTCGTGGGA 2174

Qy    251 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 270
Db   2175 ACAGTATGTGCCATTGCCAATGGGGGGCTTCAGCCGGCATTTTCAGTCATATTCTCAGAG 2234

Qy    271 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 290
Db   2235 ATCATAGCGATTTTTGGACCAGGCGATGATGCAGTGAAGCAGCAGAAGTGCAACATATTC 2294

Qy    291 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 310
Db   2295 TCTTTGATTTTCTTATTTCTGGGAATTATTTCTTTTTTTACTTTCTTCCTTCAGGGTTTC 2354

Qy    311 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 330
Db   2355 ACGTTTGGGAAAGCTGGCGAGATCCTCACCAGAAGACTGCGGTCAATGGCTTTTAAAGCA 2414

Qy    331 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 350
Db   2415 ATGCTAAGACAGGACATGAGCTGGTTTGATGACCATAAAAACAGTACTGGTGCACTTTCT 2474

Qy    351 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 370
Db   2475 ACAAGACTTGCCACAGATGCTGCCCAAGTCCAAGGAGCCACAGGAACCAGGTTGGCTTTA 2534

Qy    371 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrp 390
Db   2535 ATTGCACAGAATATAGCTAACCTTGGAACTGGTATTATCATATCATTTATCTACGGTTGG 2594

Qy    391 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 410
Db   2595 CAGTTAACCCTATTGCTATTAGCAGTTGTTCCAATTATTGCTGTGTCAGGAATTGTTGAA 2654

Qy    411 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 430
Db   2655 ATGAAATTGTTGGCTGGAAATGCCAAAAGAGATAAAAAAGAACTGGAAGCTGCTGGAAAG 2714

Qy    431 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 450
Db   2715 ATTGCAACAGAGGCAATAGAAAATATTAGGACAGTTGTGTCTTTGACCCAGGAAAGAAAA 2774

Qy    451 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 470
Db   2775 TTTGAATCAATGTATGTTGAAAAATTGTATGGACCTTACAGGAATTCTGTGCAGAAGGCA 2834

Qy    471 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 490
Db   2835 CACATCTATGGAATTACTTTTAGTATCTCACAAGCATTTATGTATTTTTCCTATGCCGGT 2894

Qy    491 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 510
Db   2895 TGTTTTCGATTTGGTGCATATCTCATTGTGAATGGACATATGCGCTTCAGAGATGTTATT 2954

Qy    511 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 530
Db   2955 CTGGTGTTTTCTGCAATTGTATTTGGTGCAGTGGCTCTAGGACATGCCAGTTCATTTGCT 3014

Qy    531 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 550
Db   3015 CCAGACTATGCTAAAGCTAAGCTGTCTGCAGCCCACTTATTCATGCTGTTTGAAAGACAA 3074

Qy    551 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 570
Db   3075 CCTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAGCCTGATAAATTTGAAGGAAATATA 3134

Qy    571 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 590
Db   3135 ACATTTAATGAAGTCGTGTTCAACTATCCCACCCGAGCAAACGTGCCAGTGCTTCAGGGG 3194

Qy    591 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 610
Db   3195 CTGAGCCTGGAGGTGAAGAAAGGCCAGACACTAGCCCTGGTGGGCAGCAGTGGCTGTGGG 3254

Qy    611 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 630
Db   3255 AAGAGCACGGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCGGGGACAGTGCTT 3314

Qy    631 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 650
Db   3315 CTCGATGGTCAAGAAGCAAAGAAACTCAATGTCCAGTGGCTCAGAGCTCAACTCGGAATC 3374

Qy    651 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 670
Db   3375 GTGTCTCAGGAGCCTATCCTATTTGACTGCAGCATTGCCGAGAATATTGCCTATGGAGAC 3434

Qy    671 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 690
```

```
Db    3435 AACAGCCGGGTTGTATCACAGGATGAAATTGTGAGTGCAGCCAAAGCTGCCAACATACAT 3494

Qy     691 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 710

Db    3495 CCTTTCATCGAGACGTTACCCCACAAATATGAAACAAGAGTGGGAGATAAGGGGACTCAG 3554

Qy     711 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 730

Db    3555 CTCTCAGGAGGTCAAAAACAGAGGATTGCTATTGCCCGAGCCCTCATCAGACAACCTCAA 3614

Qy     731 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 750

Db    3615 ATCCTCCTGTTGGATGAAGCTACATCAGCTCTGGATACTGAAAGTGAAAAGGTTGTCCAA 3674

Qy     751 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 770

Db    3675 GAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCC 3734

Qy     771 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 790

Db    3735 ACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGGAGAGTCAAGGAGCATGGC 3794

Qy     791 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 810

Db    3795 ACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3854
```

Search completed: March 31, 2003, 02:09:55
Job time : 11128.4 secs

```
                    GenCore version 5.1.4_p5_4578
          Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:          March 30, 2003, 01:20:01 ; Search time 793.594 Seconds
                                            (without alignments)
                                            2304.231 Million cell updates/sec

Title:           US-09-873-409-2
Perfect score:   4079
Sequence:        1 MVDENDIRALNVRHYRDHIG..........QELLRNRDIYFKLVNAQSVQ 812

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters:      4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115416_19216/app_query.fasta_1.7544
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_4247_@runat_27032003_115416_19216 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       N_Geneseq_101002:*
                 1:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
                 2:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
                 3:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
                 4:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
                 5:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
                 6:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
                 7:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
                 8:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
                 9:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
                 10:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
                 11:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
                 12:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
                 13:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
                 14:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
                 15:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
                 16:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
                 17:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
                 18:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
                 19:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
                 20:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
                 21:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
                 22:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
                 23:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
                 24:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

  Pred. No. is the number of results predicted by chance to have a
  score greater than or equal to the score of the result being printed,
  and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3919	96.1	3699	24	ABK83223	Human transporter
2	2465.5	60.4	3924	21	AAZ94742	Human ATP binding
3	2465.5	60.4	3924	21	AAZ88974	Human MDR-3 DNA.
4	2465.5	60.4	3924	24	ABN95801	Gene #2299 used to
5	2442	59.9	4279	22	AAD03488	Dog P-glycoprotein
6	2442	59.9	4279	22	AAD03504	Dog P-glycoprotein
7	2442	59.9	4279	22	AAD03505	Dog P-glycoprotein
8	2441	59.8	4186	22	AAF86127	Cynomologous monke
9	2441	59.8	4195	22	AAF86128	Cynomologous monke
10	2441	59.8	4279	22	AAD03506	Dog P-glycoprotein
11	2441	59.8	4369	21	AAZ52047	Rat multidrug resi
12	2441	59.8	4425	21	AAZ52048	Rat multidrug resi
13	2438	59.8	4317	22	AAD03489	Dog P-glycoprotein
14	2436	59.7	3840	24	ABL91687	Human polynucleoti
15	2436	59.7	3860	21	AAZ49332	Human wild-type mu
16	2436	59.7	3860	21	AAZ49333	Human G185V mutant
17	2436	59.7	3860	24	ABA94365	Human BCRP DNA rel
18	2436	59.7	3860	24	ABA94366	Human BCRP DNA rel
19	2436	59.7	3988	21	AAZ88973	Human MDR-1 DNA.
20	2436	59.7	4264	19	AAV66533	Mutated human P-gl
21	2436	59.7	4264	19	AAV66534	Mutated human P-gl
22	2436	59.7	4349	22	AAH57442	Human intestine ce
23	2436	59.7	4646	21	AAZ94738	Human ATP binding
24	2436	59.7	4646	24	AAD38994	Human mdr1 gene.
25	2436	59.7	4646	24	ABL68592	Kidney cancer rela
26	2436	59.7	4646	24	ABL68880	Kidney cancer rela
27	2436	59.7	4669	8	AAN70752	Sequence of human
28	2436	59.7	6505	17	AAT13394	Hybrid vector pSF-
29	2436	59.7	8630	21	AAZ24041	Retroviral M4 mdr-
30	2436	59.7	8630	21	AAZ24042	Retroviral vector
31	2433	59.6	4378	11	AAQ04522	Multidrug Resistan
32	2433	59.6	4669	14	AAQ52726	Sequence of human
33	2431.5	59.6	3912	24	ABK63653	Rat sequence diffe
34	2425	59.5	4646	15	AAQ72872	Human multidrug re
35	2425	59.5	4788	21	AAZ49335	Murine multidrug r
36	2425	59.5	4788	24	ABA94368	Mouse BCRP DNA rel
37	2413.5	59.2	4233	21	AAZ90198	Rat mdr1b2 (multis
38	2413.5	59.2	4233	22	AAF27498	Rat mdr1b2 multidr
39	2412	59.1	4669	19	AAV32645	Human P glycoprote
40	2412	59.1	4669	24	ABK52041	cDNA encoding huma
41	2404.5	58.9	4189	21	AAZ49334	Murine multidrug r
42	2404.5	58.9	4189	24	ABA94367	Mouse BCRP DNA rel
43	2404.5	58.9	4313	14	AAQ38950	Mouse multidrug re
44	2372	58.2	4254	24	ABK63517	Rat sequence diffe
45	2280	55.9	2726	15	AAQ70907	Multidrug-resistan

ALIGNMENTS

```
RESULT 1
ABK83223
ID   ABK83223 standard; cDNA; 3699 BP.
XX
AC   ABK83223;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human transporter and ion channel, TRICH14, Incyte ID 7472030CB1, cDNA.
XX
KW   Human; ss; gene; transporter and ion channel; TRICH; transport disorder;
KW   neurological disorder; muscle disorder; immunological disorder; cancer;
KW   scleroderma; systemic lupus erythematosus; allergy; leukaemia;
KW   cell proliferative disorder; cervical cancer; breast cancer;
KW   neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;
KW   myotonic dystrophy; catatonia; endocrine disorder; diabetes;
KW   Grave's disease; gastrointestinal disorder; Crohn's disease;
KW   renal disorder; Good pasture's syndrome; viral infection; cirrhosis;
```

```
KW   bacterial infection; fungal infection; parasitic infection;
KW   protozoal infection; helminthic infection; cardiovascular disorder;
KW   atherosclerosis; hepatic disease.
XX
OS   Homo sapiens.
XX
PN   WO200240541-A2.
XX
PD   23-MAY-2002.
XX
PF   25-OCT-2001; 2001WO-US46055.
XX
PR   27-OCT-2000; 2000US-243989P.
PR   03-NOV-2000; 2000US-245904P.
PR   09-NOV-2000; 2000US-247673P.
PR   17-NOV-2000; 2000US-249661P.
PR   20-NOV-2000; 2000US-252232P.
PR   01-DEC-2000; 2000US-250790P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Tang YT,  Yue H,  Nguyen DB,  Hafalia AJA,  Elliott VS,  Lu Y;
PI   Walia NK,  Yao MG,  Baughn MR,  Gandhi AR,  Ding L,  Sanjanwala M;
PI   Ramkumar J,  Arvizu C,  Gietzen KJ,  Lal PG,  Azimzai Y,  Khan FA;
PI   Thangavelu K,  Thornton M,  Lu DAM,  Tribouley CM,  Warren BA;
PI   Ison CH,  Das D,  Raumann BE,  Policky JL,  Kearney L;
XX
DR   WPI; 2002-463570/49.
DR   P-PSDB; ABG61544.
XX
PT   New transporters and ion channels (TRICH) polypeptides, useful for
PT   diagnosing, preventing, and treating disorders associated with an
PT   abnormal expression or activity of TRICH, e.g. immunological, muscular
PT   or renal disorders  -
XX
PS   Claim 5; Page 172; 178pp; English.
XX
CC   The invention relates to human transporters and ion channels (TRICH)
CC   polypeptides, a naturally occurring amino acid sequence 90 % identical to
CC   TRICH, a biologically active fragment of TRICH or an immunogenic fragment
CC   of TRICH. Also included are an isolated polynucleotide encoding TRICH,
CC   a recombinant polynucleotide comprising a promoter sequence operably
CC   linked to the TRICH polynucleotide, a cell transformed with the
CC   recombinant polynucleotide, a transgenic organism comprising the
CC   recombinant polynucleotide, an isolated antibody that binds specifically
CC   to TRICH, and screening for compounds which bind to TRICH, modulate
CC   TRICH, modulate TRICH expression or are ant/agonists of TRICH.
CC   The polypeptides are useful for diagnosing, treating, and
CC   preventing transport, neurological, muscle, immunological disorders
CC   (e.g. scleroderma, systemic lupus erythematosus, allergies), cell
CC   proliferative disorders such as cancers (e.g. leukaemia, cervical or
CC   breast cancers), neurodegenerative disorders (e.g. Parkinson's disease,
CC   Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy,
CC   catatonia), endocrine disorders (e.g. diabetes, Grave's disease),
CC   gastrointestinal disorders (e.g. Crohn's disease), renal disorders
CC   (e.g. Good pasture's syndrome), viral, bacterial, fungal, parasitic,
CC   protozoal and helminthic infections, cardiovascular disorders (e.g.
CC   atherosclerosis), or hepatic diseases (e.g. cirrhosis) and many
CC   other diseases and disorders detailed in the specification. They can also
CC   be used in assessing the effects of exogenous compounds on the
CC   expression of nucleic acid and amino acid sequences of transporters and
CC   ion channels. TRICH or its fragments may also be used in screening for
CC   compounds that specifically bind to and modulate the activity of TRICH.
CC   The polynucleotides can be used to create knock-in humanised animals or
CC   transgenic animals to model human disease. The present sequence
CC   encodes a TRICH protein.
XX
SQ   Sequence 3699 BP; 1116 A; 707 C; 860 G; 1016 T; 0 other;

Alignment Scores:
Pred. No.:              0               Length:         3699
Score:                  3919.00         Matches:        785
Percent Similarity:     96.80%          Conservative:   1
```

```
Best Local Similarity:  96.67%        Mismatches:   0
Query Match:            96.08%        Indels:       26
DB:                     24            Gaps:         1

US-09-873-409-2 (1-812) x ABK83223 (1-3699)

Qy          1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1339 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 1398

Qy         21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1399 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 1458

Qy         41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1459 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 1518

Qy         61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1519 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 1578

Qy         81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1579 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 1638

Qy        101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1639 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 1698

Qy        121 AlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThr 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1699 GCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTAGCACACCGACTTTCTACT 1758

Qy        141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1759 ATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATGCTGGCGGAGAAAGGAGCG 1818

Qy        161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1819 CATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTTGTGATGTCACAGGATATT 1878

Qy        181 LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1879 AAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCA 1938

Qy        201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSer 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1939 CTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCC 1998

Qy        221 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1999 ACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAAC 2058

Qy        241 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2059 AAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCTGTTCTAAATGGAACTGTT 2118

Qy        261 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2119 CATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATGTTTGGAAATAATGATAAA 2178

Qy        281 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2179 ACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTCGTCATTTTGGGTGTTATT 2238

Qy        301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 320
              |||||||||||||||||||||
Db       2239 TGCTTTGTCAGTTATTTCATG--------------------------------------- 2259

Qy        321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 340
                                                     |||||||||||||||||||||
Db       2260 ---------------------------------------CAGGATATTGCCTGGTTTGAT 2280
```

```
Qy     341 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2281 GAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCCATAGATATAGCACAAATT 2340

Qy     361 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2341 CAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAATGCAACTAACATGGGACTT 2400

Qy     381 SerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAla 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2401 TCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTCCTGATTCTGAGTATTGCT 2460

Qy     401 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2461 CCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATGACTGGATTTGCCAACAAA 2520

Qy     421 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2521 GATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAGAATATACGT 2580

Qy     441 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2581 ACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATGTATGAAGAGATGCTTCAG 2640

Qy     461 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2641 ACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTATGCATTCAGC 2700

Qy     481 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2701 CATGCCTTTATATATTTTGCCTATGCGGCAGGGTTTCGATTTGGAGCCTATTTAATTCAA 2760

Qy     501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2761 GCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACTGCAATTGCATATGGAGCT 2820

Qy     521 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540
           ||||||||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db    2821 ATGGCCATCGGAGAAACGCTCGTTTTGGCTCCTGAATATTCCAAAGCCAAATCGGGGGCT 2880

Qy     541 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2881 GCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGACAGCCGCAGTCAAGAAGGG 2940

Qy     561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2941 AAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTCTTCTATCCA 3000

Qy     581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3001 TGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACA 3060

Qy     601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3061 GTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCTGTTCAACTTCTGCAGAGA 3120

Qy     621 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3121 CTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTGGATGCAAAAGAATTGAAT 3180

Qy     641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3181 GTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAGCCTGTGCTCTTCAACTGC 3240

Qy     661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3241 AGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTGGTGCCATTAGATGAGATC 3300

Qy     681 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3301 AAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAAGGTCTCCCTGAGAAATAC 3360

Qy     701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3361 AACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCT 3420

Qy     721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3421 ATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCC 3480

Qy     741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3481 CTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGATAAAGCCAGGACGGGAAGG 3540

Qy     761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3541 ACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGGTT 3600

Qy     781 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3601 CTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAGCTCCTGAGAAATCGAGAC 3660

Qy     801 IleTyrPheLysLeuValAsnAlaGlnSerValGln 812
           ||||||||||||||||||||||||||||||||||||
Db    3661 ATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3696
```

```
RESULT 2
AAZ94742
ID   AAZ94742 standard; cDNA; 3924 BP.
XX
AC   AAZ94742;
XX
DT   01-AUG-2000  (first entry)
XX
DE   Human ATP binding cassette ABCB4 (MDR3) cDNA.
XX
KW   ABCB4; ATP binding cassette; human; cholesterol; lipid disorder;
KW   atherosclerosis; lipid disorder; dyslipidemia; psoriasis;
KW   lupus erythematosus; diagnosis; gene therapy; MDR3;
KW   multidrug resistance; chromosome 7q21; ss.
XX
OS   Homo sapiens.
XX
PN   WO200018912-A2.
XX
PD   06-APR-2000.
XX
PF   21-SEP-1999;   99WO-EP06991.
XX
PR   25-SEP-1998;   98US-0101706.
XX
PA   (FARB ) BAYER AG.
XX
PI   Schmitz G,  Klucken J;
XX
DR   WPI; 2000-293151/25.
XX
PT   Adenosine triphosphate binding proteins useful for identifying agents
PT   for treating atherosclerosis and other inflammatory disorders -
XX
PS   Claim 9; Page 117-119; 154pp; English.
XX
CC   The present sequence is that of human ATP binding cassette
CC   subfamily B protein ABCB4 cDNA.  The cDNA was identified using a
CC   differential display method in which monocytes from peripheral
CC   blood were subjected to macrophage differentiation and cholesterol
CC   loading with acetylated low density lipoproteins and subsequent
CC   deloading with high density lipoprotein (HDL3) to identify
CC   cholesterol sensitive genes.  The gene maps to chromosome 7q21
CC   and is also termed MDR3 (multidrug resistance).  The invention
CC   provides cholesterol-sensitive ABC genes (see AAZ94734-63).  These
CC   genes, and polypeptides encoded by them, can be used for diagnostic
CC   and therapeutic applications, and for biochemical or cell-based
CC   assays to screen for pharmacologically active modulator compounds
CC   useful for the treatment of lipid disorders, atherosclerosis or
```

```
CC   other inflammatory diseases such as psoriasis and lupus
CC   erythematosus.
XX
SQ   Sequence 3924 BP; 1145 A; 790 C; 977 G; 1012 T; 0 other;

Alignment Scores:
Pred. No.:              7.76e-229     Length:         3924
Score:                  2465.50       Matches:        478
Percent Similarity:     76.46%        Conservative:   149
Best Local Similarity:  58.29%        Mismatches:     182
Query Match:            60.44%        Indels:         11
DB:                     21            Gaps:           3

US-09-873-409-2 (1-812) x AAZ94742 (1-3924)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db   1395 ATTGATGGGCAGGATATTAGGAACTTTAATGTAAACTATCTGAGGGAAATCATTGGTGTG 1454

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db   1455 GTGAGTCAGGAGCCGGTGCTGTTTTCCACCACAATTGCTGAAAATATTTGTTATGGCCGT 1514

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db   1515 GGAAATGTAACCATGGATGAGATAAAGAAAGCTGTCAAAGAGGCCAACGCCTATGAGTTT 1574

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db   1575 ATCATGAAATTACCACAGAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGCTGAGT 1634

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db   1635 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTT 1694

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db   1695 CTGCTGGATGAGGCCACGTCAGCATTGGACACAGAAAGTGAAGCTGAGGTACAGGCAGCT 1754

Qy    122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db   1755 CTGGATAAGGCCAGAGAAGGCCGGACCACCATTGTGATAGCACACCGACTGTCTACGGTC 1814

Qy    142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db   1815 CGAAATGCAGATGTCATCGCTGGGTTTGAGGATGGAGTAATTGTGGAGCAAGGAAGCCAC 1874

Qy    162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal------------------ 175
Db   1875 AGCGAACTGATGAAGAAGGAAGGGGTGTACTTCAAACTTGTCAACATGCAGACATCAGGA 1934

Qy    176 ---------MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyr 192
Db   1935 AGCCAGATCCAGTCAGAAGAATTTGAACTAAATGATGAAAAGGCTGCCACTAGAATGGCC 1994

Qy    193 SerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys---Ser 211
Db   1995 CCAAATGGCTGGAAATCTCGCCTATTTAGGCATTCTACTCAGAAAAACCTTAAAAATTCA 2054

Qy    212 AspPheIleAspLysAla---GluGluSerThrGlnSerLysGluIleSerLeuProGlu 230
Db   2055 CAAATGTGTCAGAAGAGCCTTGATGTGGAAACCGATGGACTTGAAGCAAATGTGCCACCA 2114

Qy    231 ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGly 250
Db   2115 GTGTCCTTTCTGAAGGTCCTGAAACTGAATAAAACAGAATGGCCCTACTTTGTCGTGGGA 2174

Qy    251 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 270
Db   2175 ACAGTATGTGCCATTGCCAATGGGGGGCTTCAGCCGGCATTTTCAGTCATATTCTCAGAG 2234

Qy    271 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 290
Db   2235 ATCATAGCGATTTTTGGACCAGGCGATGATGCAGTGAAGCAGCAGAAGTGCAACATATTC 2294

Qy    291 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 310
Db   2295 TCTTTGATTTTCTTATTTCTGGGAATTATTTCTTTTTTACTTTCTTCCTTCAGGGTTTC 2354

Qy    311 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 330
Db   2355 ACGTTTGGGAAAGCTGGCGAGATCCTCACCAGAAGACTGCGGTCAATGGCTTTTAAAGCA 2414

Qy    331 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 350
Db   2415 ATGCTAAGACAGGACATGAGCTGGTTTGATGACCATAAAAACAGTACTGGTGCACTTTCT 2474

Qy    351 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 370
Db   2475 ACAAGACTTGCCACAGATGCTGCCCAAGTCCAAGGAGCCACAGGAACCAGGTTGGCTTTA 2534

Qy    371 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrp 390
Db   2535 ATTGCACAGAATATAGCTAACCTTGGAACTGGTATTATCATATCATTTATCTACGGTTGG 2594

Qy    391 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 410
Db   2595 CAGTTAACCCTATTGCTATTAGCAGTTGTTCCAATTATTGCTGTGTCAGGAATTGTTGAA 2654

Qy    411 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 430
Db   2655 ATGAAATTGTTGGCTGGAAATGCCAAAAGAGATAAAAAAGAACTGGAAGCTGCTGGAAAG 2714

Qy    431 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 450
Db   2715 ATTGCAACAGAGGCAATAGAAAATATTAGGACAGTTGTGTCTTTGACCCAGGAAAGAAAA 2774

Qy    451 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 470
Db   2775 TTTGAATCAATGTATGTTGAAAAATTGTATGGACCTTACAGGAATTCTGTGCAGAAGGCA 2834

Qy    471 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 490
Db   2835 CACATCTATGGAATTACTTTTAGTATCTCACAAGCATTTATGTATTTTTCCTATGCCGGT 2894

Qy    491 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 510
Db   2895 TGTTTTCGATTTGGTGCATATCTCATTGTGAATGGACATATGCGCTTCAGAGATGTTATT 2954

Qy    511 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 530
Db   2955 CTGGTGTTTTCTGCAATTGTATTTGGTGCAGTGGCTCTAGGACATGCCAGTTCATTTGCT 3014

Qy    531 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 550
Db   3015 CCAGACTATGCTAAAGCTAAGCTGTCTGCAGCCCACTTATTCATGCTGTTTGAAAGACAA 3074

Qy    551 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 570
Db   3075 CCTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAGCCTGATAAATTTGAAGGAAATATA 3134

Qy    571 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 590
Db   3135 ACATTTAATGAAGTCGTGTTCAACTATCCCACCCGAGCAAACGTGCCAGTGCTTCAGGGG 3194

Qy    591 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 610
Db   3195 CTGAGCCTGGAGGTGAAGAAAGGCCAGACACTAGCCCTGGTGGGCAGCAGTGGCTGTGGG 3254

Qy    611 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 630
Db   3255 AAGAGCACGGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCGGGGACAGTGCTT 3314

Qy    631 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 650
Db   3315 CTCGATGGTCAAGAAGCAAAGAAACTCAATGTCCAGTGGCTCAGAGCTCAACTCGGAATC 3374
```

```
Qy     651 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 670
Db    3375 GTGTCTCAGGAGCCTATCCTATTTGACTGCAGCATTGCCGAGAATATTGCCTATGGAGAC 3434

Qy     671 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 690
Db    3435 AACAGCCGGGTTGTATCACAGGATGAAATTGTGAGTGCAGCCAAAGCTGCCAACATACAT 3494

Qy     691 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 710
Db    3495 CCTTTCATCGAGACGTTACCCCACAAATATGAAACAAGAGTGGGAGATAAGGGGACTCAG 3554

Qy     711 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 730
Db    3555 CTCTCAGGAGGTCAAAAACAGAGGATTGCTATTGCCCGAGCCCTCATCAGACAACCTCAA 3614

Qy     731 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 750
Db    3615 ATCCTCCTGTTGGATGAAGCTACATCAGCTCTGGATACTGAAAGTGAAAAGGTTGTCCAA 3674

Qy     751 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 770
Db    3675 GAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCC 3734

Qy     771 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 790
Db    3735 ACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGGAGAGTCAAGGAGCATGGC 3794

Qy     791 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 810
Db    3795 ACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3854
```

```
RESULT 3
AAZ88974
ID   AAZ88974 standard; DNA; 3924 BP.
XX
AC   AAZ88974;
XX
DT   30-MAY-2000  (first entry)
XX
DE   Human MDR-3 DNA.
XX
KW   Phosphatidylcholine; treatment; disease; multidrug resistance; MDR-3;
KW   stomach; colon; ulcerative colitis; pouchitis; large intestine; human;
KW   inflammation; Crohn's disease; diverse colitis; carcinoma; gene therapy;
KW   ileum; anti-ulcer; anti-inflammatory; cytostatic; ss.
XX
OS   Homo sapiens.
XX
PN   WO200007577-A2.
XX
PD   17-FEB-2000.
XX
PF   06-AUG-1999;   99WO-DE02426.
XX
PR   06-AUG-1998;   98DE-1035526.
PR   15-DEC-1998;   98DE-1057750.
XX
PA   (STRE/) STREMMEL W.
XX
PI   Stremmel W;
XX
DR   WPI; 2000-195439/17.
XX
PT   Using phosphatidylcholine having mucous membrane protective activity to
PT   prevent or treat large intestinal diseases, such as ulcerative colitis
PT   or pouchitis  -
XX
PS   Disclosure; Page 17-22; 22pp; German.
XX
CC   This invention describes a novel medicament containing a therapeutically
CC   effective amount of phosphatidylcholine to treat diseases, where
CC   phosphatidylcholine has an advantageous mucous membrane protective
CC   activity in the large intestine. The invention describes (1) a method
CC   to regulate presence or absence of MDR(multidrug resistance)-3
CC   transcription in a sample from patients with suspicion of stomach and
CC   colon diseases, where treatment with phosphatidylcholine is indicated;
CC   (2) a diagnostic test, to detect ulcerative colitis, pouchitis, large
CC   intestinal inflammation, Crohn's disease, diverse colitis, infectious
CC   enteritis/colitis, inflammation through X-ray treatment, antibiotics,
CC   chemotherapeutics, drugs or chemicals or large intestinal carcinoma,
CC   contains MDR3-specific primers to detect mutations in the MDR3 gene
CC   family and their interactions partners (in particular transcription
CC   factors); and (3) a method to manufacture a medicament for local gene
CC   therapy in the ileum with the ideal MDR3-analogous genes for patients
CC   with missing or decreased expression of MDR3-analogous proteins. The
CC   products of the invention have anti-ulcer, anti-inflammatory and
CC   cytostatic activity. The orally administered delayed-release form of
CC   phosphatidylcholine prevents premature resorption and provides for
CC   targeted release in the lower section of the small or large intestine.
CC   This sequence encodes the human MDR-3 protein described in the method of
CC   the invention.
XX
SQ   Sequence 3924 BP; 1145 A; 790 C; 977 G; 1012 T; 0 other;

Alignment Scores:
Pred. No.:              7.76e-229      Length:         3924
Score:                  2465.50        Matches:        478
Percent Similarity:     76.46%         Conservative:   149
Best Local Similarity:  58.29%         Mismatches:     182
Query Match:            60.44%         Indels:         11
DB:                     21             Gaps:           3

US-09-873-409-2 (1-812) x AAZ88974 (1-3924)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db    1395 ATTGATGGGCAGGATATTAGGAACTTTAATGTAAACTATCTGAGGGAAATCATTGGTGTG 1454

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db    1455 GTGAGTCAGGAGCCGGTGCTGTTTTCCACCACAATTGCTGAAAATATTTGTTATGGCCGT 1514

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db    1515 GGAAATGTAACCATGGATGAGATAAAGAAAGCTGTCAAAGAGGCCAACGCCTATGAGTTT 1574

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db    1575 ATCATGAAATTACCACAGAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGCTGAGT 1634

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db    1635 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTT 1694

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db    1695 CTGCTGGATGAGGCCACGTCAGCATTGGACACAGAAAGTGAAGCTGAGGTACAGGCAGCT 1754

Qy     122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db    1755 CTGGATAAGGCCAGAGAAGGCCGGACCACCATTGTGATAGCACACCGACTGTCTACGGTC 1814

Qy     142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db    1815 CGAAATGCAGATGTCATCGCTGGGTTTGAGGATGGAGTAATTGTGGAGCAAGGAAGCCAC 1874

Qy     162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal------------------ 175
Db    1875 AGCGAACTGATGAAGAAGGAAGGGGTGTACTTCAAACTTGTCAACATGCAGACATCAGGA 1934

Qy     176 ---------MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyr 192
Db    1935 AGCCAGATCCAGTCAGAAGAATTTGAACTAAATGATGAAAAGGCTGCCACTAGAATGGCC 1994
```

```
Qy    193 SerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys---Ser 211
Db   1995 CCAAATGGCTGGAAATCTCGCCTATTTAGGCATTCTACTCAGAAAAACCTTAAAAATTCA 2054

Qy    212 AspPheIleAspLysAla---GluGluSerThrGlnSerLysGluIleSerLeuProGlu 230
Db   2055 CAAATGTGTCAGAAGAGCCTTGATGTGGAAACCGATGGACTTGAAGCAAATGTGCCACCA 2114

Qy    231 ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGly 250
Db   2115 GTGTCCTTTCTGAAGGTCCTGAAACTGAATAAAACAGAATGGCCCTACTTTGTCGTGGGA 2174

Qy    251 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 270
Db   2175 ACAGTATGTGCCATTGCCAATGGGGGGCTTCAGCCGGCATTTTCAGTCATATTCTCAGAG 2234

Qy    271 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 290
Db   2235 ATCATAGCGATTTTTGGACCAGGCGATGATGCAGTGAAGCAGCAGAAGTGCAACATATTC 2294

Qy    291 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 310
Db   2295 TCTTTGATTTCTTATTTCTGGGAATTATTTCTTTTTTACTTTCTTCCTTCAGGGTTTC 2354

Qy    311 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 330
Db   2355 ACGTTTGGGAAAGCTGGCGAGATCCTCACCAGAAGACTGCGGTCAATGGCTTTTAAAGCA 2414

Qy    331 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 350
Db   2415 ATGCTAAGACAGGACATGAGCTGGTTTGATGACCATAAAAACAGTACTGGTGCACTTTCT 2474

Qy    351 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 370
Db   2475 ACAAGACTTGCCACAGATGCTGCCCAAGTCCAAGGAGCCACAGGAACCAGGTTGGCTTTA 2534

Qy    371 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrp 390
Db   2535 ATTGCACAGAATATAGCTAACCTTGGAACTGGTATTATCATATCATTTATCTACGGTTGG 2594

Qy    391 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 410
Db   2595 CAGTTAACCCTATTGCTATTAGCAGTTGTTCCAATTATTGCTGTGTCAGGAATTGTTGAA 2654

Qy    411 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 430
Db   2655 ATGAAATTGTTGGCTGGAAATGCCAAAAGAGATAAAAAAGAACTGGAAGCTGCTGGAAAG 2714

Qy    431 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 450
Db   2715 ATTGCAACAGAGGCAATAGAAAATATTAGGACAGTTGTGTCTTTGACCCAGGAAAGAAAA 2774

Qy    451 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 470
Db   2775 TTTGAATCAATGTATGTTGAAAAATTGTATGGACCTTACAGGAATTCTGTGCAGAAGGCA 2834

Qy    471 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 490
Db   2835 CACATCTATGGAATTACTTTTAGTATCTCACAAGCATTTATGTATTTTTCCTATGCCGGT 2894

Qy    491 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 510
Db   2895 TGTTTTCGATTTGGTGCATATCTCATTGTGAATGGACATATGCGCTTCAGAGATGTTATT 2954

Qy    511 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 530
Db   2955 CTGGTGTTTTCTGCAATTGTATTTGGTGCAGTGGCTCTAGGACATGCCAGTTCATTTGCT 3014

Qy    531 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 550
Db   3015 CCAGACTATGCTAAAGCTAAGCTGTCTGCAGCCCACTTATTCATGCTGTTTGAAAGACAA 3074

Qy    551 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 570
Db   3075 CCTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAGCCTGATAAATTTGAAGGAAATATA 3134

Qy    571 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 590
Db   3135 ACATTTAATGAAGTCGTGTTCAACTATCCCACCCGAGCAAACGTGCCAGTGCTTCAGGGG 3194

Qy    591 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 610
Db   3195 CTGAGCCTGGAGGTGAAGAAAGGCCAGACACTAGCCCTGGTGGGCAGCAGTGGCTGTGGG 3254

Qy    611 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 630
Db   3255 AAGAGCACGGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCGGGGACAGTGCTT 3314

Qy    631 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 650
Db   3315 CTCGATGGTCAAGAAGCAAAGAAACTCAATGTCCAGTGGCTCAGAGCTCAACTCGGAATC 3374

Qy    651 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 670
Db   3375 GTGTCTCAGGAGCCTATCCTATTTGACTGCAGCATTGCCGAGAATATTGCCTATGGAGAC 3434

Qy    671 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 690
Db   3435 AACAGCCGGGTTGTATCACAGGATGAAATTGTGAGTGCAGCCAAAGCTGCCAACATACAT 3494

Qy    691 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 710
Db   3495 CCTTTCATCGAGACGTTACCCCACAAATATGAAACAAGAGTGGGAGATAAGGGGACTCAG 3554

Qy    711 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 730
Db   3555 CTCTCAGGAGGTCAAAAACAGAGGATTGCTATTGCCCGAGCCCTCATCAGACAACCTCAA 3614

Qy    731 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 750
Db   3615 ATCCTCCTGTTGGATGAAGCTACATCAGCTCTGGATACTGAAAGTGAAAAGGTTGTCCAA 3674

Qy    751 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 770
Db   3675 GAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCC 3734

Qy    771 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 790
Db   3735 ACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGGAGAGTCAAGGAGCATGGC 3794

Qy    791 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 810
Db   3795 ACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3854
```

```
RESULT 4
ABN95801
ID   ABN95801 standard; DNA; 3924 BP.
XX
AC   ABN95801;
XX
DT   13-AUG-2002  (first entry)
XX
DE   Gene #2299 used to diagnose liver cancer.
XX
KW   Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW   metastatic liver tumour; cytostatic; expression profile; disease state;
KW   disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS   Homo sapiens.
XX
PN   WO200229103-A2.
XX
PD   11-APR-2002.
XX
PF   02-OCT-2001; 2001WO-US30589.
XX
```

```
PR   02-OCT-2000; 2000US-237054P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Horne D,  Alvares C,  Peres-Da-Silva S,  Vockley JG;
XX
DR   WPI; 2002-426119/45.
XX
PT   Diagnosing and detecting the progression of liver cancer,
PT   hepatocellular carcinoma or metastatic liver tumor in a patient,
PT   involves detecting the level of expression of two or more genes in a
PT   liver tissue sample  -
XX
PS   Claim 1; SEQ ID NO 2299; 298pp; English.
XX
CC   The invention relates to a novel method for diagnosing and detecting the
CC   progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC   tumour in a patient, and differentiating metastatic liver cancer from
CC   hepatocellular carcinoma in a patient, involving detecting the level of
CC   expression of two or more genes represented in ABN93503-ABN97455 in a
CC   tissue sample. The method of the invention has hepatotropic, and
CC   cytostatic activity. The method is useful for diagnosing and detecting
CC   the progression of liver cancer, hepatocellular carcinoma and metastatic
CC   liver carcinoma in a patient. The method is useful for identifying
CC   expression profiles which serve as useful diagnostic markers as well as
CC   markers that can be used to monitor disease states, disease progression,
CC   drug toxicity, drug efficacy and drug metabolism.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 3924 BP; 1145 A; 790 C; 977 G; 1012 T; 0 other;

Alignment Scores:
Pred. No.:              7.76e-229     Length:         3924
Score:                  2465.50       Matches:        478
Percent Similarity:     76.46%        Conservative:   149
Best Local Similarity:  58.29%        Mismatches:     182
Query Match:            60.44%        Indels:         11
DB:                     24            Gaps:           3

US-09-873-409-2 (1-812) x ABN95801 (1-3924)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
             :::|||      ||||||||      |||||   :::   |||:::   ||||||||
Db    1395 ATTGATGGGCAGGATATTAGGAACTTTAATGTAAACTATCTGAGGGAAATCATTGGTGTG 1454

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           ||||||||||||||||||||||||   |||||||||:::   ||||||   ||||||||
Db    1455 GTGAGTCAGGAGCCGGTGCTGTTTTCCACCACAATTGCTGAAAATATTTGTTATGGCCGT 1514

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
              :::||||||   :::|||:::::::::|||   :::|||||||||||||||:::|||
Db    1515 GGAAATGTAACCATGGATGAGATAAAGAAAGCTGTCAAAGAGGCCAACGCCTATGAGTTT 1574

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           ||||||:::      |||   ||||||:::||||||||||||||||:::||||||:::|||
Db    1575 ATCATGAAATTACCACAGAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGCTGAGT 1634

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1635 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTT 1694

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
           :::||||||||||||||||||||||||||||:::||||||::::::   |||||||||||
Db    1695 CTGCTGGATGAGGCCACGTCAGCATTGGACACAGAAAGTGAAGCTGAGGTACAGGCAGCT 1754

Qy     122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
           |||:::||||||   :::|||||||||||||||||:::||||||||||||||||||:::
Db    1755 CTGGATAAGGCCAGAGAAGGCCGGACCACCATTGTGATAGCACACCGACTGTCTACGGTC 1814

Qy     142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
           |||:::||||||:::|||          :::||||||::::::   |||:::|||:::|||
Db    1815 CGAAATGCAGATGTCATCGCTGGGTTTGAGGATGGAGTAATTGTGGAGCAAGGAAGCCAC 1874

Qy     162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal------------------ 175
           :::|||||||||   |||   |||:::|||:::   ||||||
Db    1875 AGCGAACTGATGAAGAAGGAAGGGGTGTACTTCAAACTTGTCAACATGCAGACATCAGGA 1934

Qy     176 ---------MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyr 192
                    |||::::::   :::      ||||||:::         :::
Db    1935 AGCCAGATCCAGTCAGAAGAATTTGAACTAAATGATGAAAAGGCTGCCACTAGAATGGCC 1994

Qy     193 SerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys---Ser 211
                       |||:::            :::   |||::::::|||   |||
Db    1995 CCAAATGGCTGGAAATCTCGCCTATTTAGGCATTCTACTCAGAAAAACCTTAAAAATTCA 2054

Qy     212 AspPheIleAspLysAla---GluGluSerThrGlnSerLysGluIleSerLeuProGlu 230
                       |||:::   :::      |||         |||   ::::::|||
Db    2055 CAAATGTGTCAGAAGAGCCTTGATGTGGAAACCGATGGACTTGAAGCAAATGTGCCACCA 2114

Qy     231 ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGly 250
           ||||||   ||||||:::|||||||||||||||   ||||||||:::   |||:::|||
Db    2115 GTGTCCTTTCTGAAGGTCCTGAAACTGAATAAAACAGAATGGCCCTACTTTGTCGTGGGA 2174

Qy     251 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 270
           |||:::   ::::::   ||||||   :::   |||   ||||||:::||||||::::::
Db    2175 ACAGTATGTGCCATTGCCAATGGGGGGCTTCAGCCGGCATTTTCAGTCATATTCTCAGAG 2234

Qy     271 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 290
           ||||||   :::|||||||||         |||             :::   |||:::
Db    2235 ATCATAGCGATTTTTGGACCAGGCGATGATGCAGTGAAGCAGCAGAAGTGCAACATATTC 2294

Qy     291 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 310
           |||:::||||||:::   ||||||:::|||   |||   ::::::|||:::||||||
Db    2295 TCTTTGATTTTCTTATTTCTGGGAATTATTTCTTTTTTACTTTCTTCCTTCAGGGTTTC 2354

Qy     311 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 330
              :::|||:::||||||||||||||||||   ||||||||   :::||||||||||||
Db    2355 ACGTTTGGGAAAGCTGGCGAGATCCTCACCAGAAGACTGCGGTCAATGGCTTTTAAAGCA 2414

Qy     331 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 350
           ||||||   ||||||::::::|||||||||:::   :::||||||||||||   |||:::
Db    2415 ATGCTAAGACAGGACATGAGCTGGTTTGATGACCATAAAAACAGTACTGGTGCACTTTCT 2474

Qy     351 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 370
           |||   ||||||   |||   ||||||:::||||||||||||:::|||:::   :::
Db    2475 ACAAGACTTGCCACAGATGCTGCCCAAGTCCAAGGAGCCACAGGAACCAGGTTGGCTTTA 2534

Qy     371 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrp 390
           :::   ||||||      |||:::|||      :::||||||||||||||||||||||||
Db    2535 ATTGCACAGAATATAGCTAACCTTGGAACTGGTATTATCATATCATTTATCTACGGTTGG 2594

Qy     391 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 410
           ::::::|||   |||:::|||::::::   |||:::::::||||||:::|||::::::|||
Db    2595 CAGTTAACCCTATTGCTATTAGCAGTTGTTCCAATTATTGCTGTGTCAGGAATTGTTGAA 2654

Qy     411 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 430
                       :::   |||   |||   :::||||||:::||||||:::   ||||||||
Db    2655 ATGAAATTGTTGGCTGGAAATGCCAAAAGAGATAAAAAAGAACTGGAAGCTGCTGGAAAG 2714

Qy     431 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 450
           ||||||||||||||||:::||||||||||||||||:::|||||||||||:::|||:::
Db    2715 ATTGCAACAGAGGCAATAGAAAATATTAGGACAGTTGTGTCTTTGACCCAGGAAAGAAAA 2774

Qy     451 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 470
           ||||||   ||||||   |||   |||            :::||||||:::   :::||||||
Db    2775 TTTGAATCAATGTATGTTGAAAAAATTGTATGGACCTTACAGGAATTCTGTGCAGAAGGCA 2834

Qy     471 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 490
              |||   |||      ::::::   |||   ||||||:::||||||:::||||||
Db    2835 CACATCTATGGAATTACTTTAGTATCTCACAAGCATTTATGTATTTTTCCTATGCCGGT 2894

Qy     491 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 510
              ||||||||||||||||||||||||         |||   |||            :::
Db    2895 TGTTTTCGATTTGGTGCATATCTCATTGTGAATGGACATATGCGCTTCAGAGATGTTATT 2954
```

```
Qy     511 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 530
           :::||||||:::||||||   :::||||||:::||||:::|||              |||
Db    2955 CTGGTGTTTTCTGCAATTGTATTTGGTGCAGTGGCTCTAGGACATGCCAGTTCATTTGCT 3014

Qy     531 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 550
           |||:::|||:::|||||||||      ||||||||||||||   |||   |||::::::
Db    3015 CCAGACTATGCTAAAGCTAAGCTGTCTGCAGCCCACTTATTCATGCTGTTTGAAAGACAA 3074

Qy     551 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 570
           |||   |||||||||   |||:::||||||   |||||||||         |||||||||:::
Db    3075 CCTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAGCCTGATAAATTTGAAGGAAATATA 3134

Qy     571 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 590
              |||   ||||||   |||   ||||||   |||     :::|||:::|||
Db    3135 ACATTTAATGAAGTCGTGTTCAACTATCCCACCCGAGCAAACGTGCCAGTGCTTCAGGGG 3194

Qy     591 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 610
           |||||||||   ::::::::|||:::|||:::|||   ||||||||||||||||||||
Db    3195 CTGAGCCTGGAGGTGAAGAAAGGCCAGACACTAGCCCTGGTGGGCAGCAGTGGCTGTGGG 3254

Qy     611 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 630
           |||||||||   ||||||||||||:::|||   |||||||||:::   |||   ||||||
Db    3255 AAGAGCACGGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCGGGGACAGTGCTT 3314

Qy     631 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 650
              ||||||   :::||||||:::|||||||||||||||||||||:::|||:::   |||
Db    3315 CTCGATGGTCAAGAAGCAAAGAAACTCAATGTCCAGTGGCTCAGAGCTCAACTCGGAATC 3374

Qy     651 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 670
           |||   |||||||||:::||||||:::|||||||||||||||||||||||||||||||||
Db    3375 GTGTCTCAGGAGCCTATCCTATTTGACTGCAGCATTGCCGAGAATATTGCCTATGGAGAC 3434

Qy     671 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 690
           |||||||||||||||      |||||||||         ||||||   |||||||||||||||
Db    3435 AACAGCCGGGTTGTATCACAGGATGAAATTGTGAGTGCAGCCAAAGCTGCCAACATACAT 3494

Qy     691 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 710
              |||||||||   ||||||   ||||||   |||:::||||||   ||||||   |||
Db    3495 CCTTTCATCGAGACGTTACCCCACAAATATGAAACAAGAGTGGGAGATAAGGGGACTCAG 3554

Qy     711 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 730
           |||||||||||||||||||||||:::||||||||||||||||||:::::::::|||:::
Db    3555 CTCTCAGGAGGTCAAAAACAGAGGATTGCTATTGCCCGAGCCCTCATCAGACAACCTCAA 3614

Qy     731 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 750
           ||||||||||||||||||||||||||||||||||||   :::||||||||||||||||||
Db    3615 ATCCTCCTGTTGGATGAAGCTACATCAGCTCTGGATACTGAAAGTGAAAAGGTTGTCCAA 3674

Qy     751 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 770
              |||||||||||||||||||   ||||||||||||:::|||:::   ||||||||||||
Db    3675 GAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCC 3734

Qy     771 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 790
              ||||||||||||||||||||||||||      ||||||::::::||||||   |||
Db    3735 ACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGGAGAGTCAAGGAGCATGGC 3794

Qy     791 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 810
           |||||||||:::||||||      :::   |||||||||   :::|||:::   |||:::
Db    3795 ACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3854
```

RESULT 5
AAD03488

```
ID   AAD03488 standard; cDNA; 4279 BP.
XX
AC   AAD03488;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) cDNA #1.
XX
KW   Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag= a
FT                   /product= "Dog P-glycoprotein (PGP) #1"
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00303.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 3; Page 58-63; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) cDNA. This
CC   sequence is also referred as Genotype C cDNA. The
CC   PGP enzyme functions as an efflux pump exporting small molecules
CC   across the cell membrane. This enzyme is a member of the ABC
CC   transporter family.
XX
SQ   Sequence 4279 BP; 1294 A; 834 C; 1008 G; 1143 T; 0 other;
```

```
Alignment Scores:
Pred. No.:              1.69e-226       Length:         4279
Score:                  2442.00         Matches:        469
Percent Similarity:     76.24%          Conservative:   160
Best Local Similarity:  56.85%          Mismatches:     178
Query Match:            59.87%          Indels:         18
DB:                     22              Gaps:           4
```

US-09-873-409-2 (1-812) x AAD03488 (1-4279)

```
Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
           :::|||      ||||||||||   :::|||||||||||||   |||:::      ||||||
Db    1379 ATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTG 1438

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           ||||||||||||||||||||||||   ||||||||||:::   ||||||:::|||||||||
Db    1439 GTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGC 1498

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
           ::::::||||||   :::|||:::|||:::|||      :::|||||||||||||||||||
Db    1499 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTT 1558

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
```

```
Db   1559 ATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGT 1618

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db   1619 GGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTT 1678

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db   1679 CTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCC 1738

Qy    122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db   1739 CTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCTCATCGTTTGTCTACAGTT 1798

Qy    142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db   1799 CGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 1858

Qy    162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLys 181
Db   1859 GATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTCACAATGCAG---ACAAGA 1915

Qy    182 LysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu 201
Db   1916 GGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAAAGTGAAAGTGATGCCTTG 1975

Qy    202 ProLeuHisSerValLysSerIleLysSerAspPheIle--------------------- 214
Db   1976 GAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAAAGAAGATCAACTCGCAGG 2032

Qy    215 -----------------------AspLysAlaGluGluSerThrGlnSerLysGluIle 226
Db   2033 AGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACAAAAGAGGACTTGAATGAG 2092

Qy    227 SerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPhe 246
Db   2093 AATGTACCTCCAGTTTCCTTCTGGAGGATTCTGAAGCTGAACTCAACTGAATGGCCTTAT 2152

Qy    247 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 266
Db   2153 TTTGTGGTTGGTATATTTTGTGCTATTATAAACGGAGGCCTGCAACCAGCATTTTCAATA 2212

Qy    267 IlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHis 285
Db   2213 ATATTTTCAAGGATTATAGGGATCTTTACCCGAGATGAGGATCCTGAAACAAAACGACAG 2272

Qy    286 AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr 305
Db   2273 AATAGTAACATGTTTTCTGTATTGTTTCTAGTCCTTGGAATTATTTCTTTTATTACATTT 2332

Qy    306 PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHis 325
Db   2333 TTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTCACTAAGCGGCTTCGATAC 2392

Qy    326 LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSer 345
Db   2393 ATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTTGATGACCCTAAAAACACC 2452

Qy    346 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365
Db   2453 ACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAAGTTAAAGGGGCTATAGGT 2512

Qy    366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385
Db   2513 TCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGGACAGGCATTATTATATCC 2572

Qy    386 PheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405
Db   2573 TTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATTGTACCCATCATTGCAATA 2632

Qy    406 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425
Db   2633 GCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAGCTA 2692

Qy    426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445
Db   2693 GAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTCCGAACTGTTGTTTCTTTG 2752

Qy    446 ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn 465
Db   2753 ACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTGCAAGTACCATACAGAAAC 2812

Qy    466 ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr 485
Db   2813 TCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATCACCCAGGCAATGATGTAT 2872

Qy    486 PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr 505
Db   2873 TTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTGGCAAATGAGTTCATGAAC 2932

Qy    506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys 525
Db   2933 TTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAG 2992

Qy    526 ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAla 545
Db   2993 GTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCAGCAGCCCACGTCATCATG 3052

Qy    546 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 565
Db   3053 ATCATTGAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCACGGCCTCAAGCCAAATACG 3112

Qy    566 CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspVal 585
Db   3113 TTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTATCCCACTCGACCAGACATC 3172

Qy    586 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605
Db   3173 CCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAGACGCTGGCCCTCGTAGGT 3232

Qy    606 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProVal 625
Db   3233 AGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAGCGCTTCTATGACCCCTTG 3292

Qy    626 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArg 645
Db   3293 GCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTGAATGTCCAGTGGCTCCGA 3352

Qy    646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665
Db   3353 GCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGACTGCAGCATTGCCGAGAAC 3412

Qy    666 IleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn 685
Db   3413 ATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAGATTATGCAGGCAGCCAAG 3472

Qy    686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 705
Db   3473 GAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAATACAACACCAGAGTAGGA 3532

Qy    706 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725
Db   3533 GACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATTGCCATAGCTCGCGCTCTT 3592

Qy    726 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745
Db   3593 GTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAGT 3652

Qy    746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765
Db   3653 GAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATC 3712

Qy    766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785
Db   3713 GCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTGGTGTTTCAGAATGGCAAA 3772
```

```
Qy    786 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 805
          :::||||||   |||||||||||:::||||||        :::   ||||||||    :::
Db   3773 GTCAAGGAGCATGGCACACATCAACAGCTGCTGGCCCAGAAAGGCATCTATTTTTCCATG 3832

Qy    806 ValAsnAlaGlnSer 810
          |||:::   |||:::
Db   3833 GTCAGTGTCCAGGCT 3847
```

```
RESULT 6
AAD03504
ID   AAD03504 standard; cDNA; 4279 BP.
XX
AC   AAD03504;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) allelic variant (Genotype A) cDNA.
XX
KW   Dog; P-glycoprotein allelic variant; multidrug transporter; MDR1;
KW   drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag=  a
FT                   /product= "Dog P-glycoprotein (PGP) allelic variant
FT                   (Genotype A) protein"
FT   allele          replace (607, C)
FT                   /*tag=  b
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00308.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 9; Page 85-90; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) allelic variant
CC   (Genotype A) cDNA. The PGP enzyme functions as an efflux pump
CC   exporting small molecules across the cell membrane. This enzyme
CC   is a member of the ABC transporter family.
XX
SQ   Sequence 4279 BP; 1295 A; 833 C; 1008 G; 1143 T; 0 other;
```

```
Alignment Scores:
Pred. No.:               1.69e-226      Length:          4279
Score:                   2442.00        Matches:         469
Percent Similarity:      76.24%         Conservative:    160
Best Local Similarity:   56.85%         Mismatches:      178
Query Match:             59.87%         Indels:          18
DB:                      22             Gaps:            4

US-09-873-409-2 (1-812) x AAD03504 (1-4279)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
          :::|||      |||||||||   :::|||||||||||||   |||:::      ||||||
Db   1379 ATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTG 1438

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
          ||||||||||||||||||||||||   |||||||||:::   ||||||:::||||||||
Db   1439 GTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGC 1498

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
          ::::::||||||   :::|||:::|||:::|||   :::||||||||||||||||||||||
Db   1499 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTT 1558

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
          ||||||:::   ||||||||||||:::|||||||||||||||:::|||||||||:::|||
Db   1559 ATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGT 1618

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1619 GGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTT 1678

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
          :::|||||||||||||||||||||||||||:::||||||::::::   ||||||   |||
Db   1679 CTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCC 1738

Qy    122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
          |||:::||||||   |||||||||||||||||||||:::||||||||||||||||||:::
Db   1739 CTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCTCATCGTTTGTCTACAGTT 1798

Qy    142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
          |||:::||||||:::|||            ||||||::::::   ||||||||   |||
Db   1799 CGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 1858

Qy    162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLys 181
             |||||||||   ::::::|||:::|||:::   ||||||      |||      :::
Db   1859 GATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTCACAATGCAG---ACAAGA 1915

Qy    182 LysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu 201
             :::   ::::::|||:::   |||      ::::::   ::::::::::::|||
Db   1916 GGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAAAGTGAAAGTGATGCCTTG 1975

Qy    202 ProLeuHisSerValLysSerIleLysSerAspPheIle--------------------- 214
             :::   |||   |||         |||      |||
Db   1976 GAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAAAGAAGATCAACTCGCAGG 2032

Qy    215 ------------------------AspLysAlaGluGluSerThrGlnSerLysGluIle 226
                                  |||:::         :::   :::
Db   2033 AGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACAAAAGAGGACTTGAATGAG 2092

Qy    227 SerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPhe 246
          ::::::|||   ||||||      :::||||||||||||||||      |||||||||:::
Db   2093 AATGTACCTCCAGTTTCCTTCTGGAGGATTCTGAAGCTGAACTCAACTGAATGGCCTTAT 2152

Qy    247 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 266
             |||:::|||         :::::::::||||||   :::   |||   |||||||||
Db   2153 TTTGTGGTTGGTATATTTTGTGCTATTATAAACGGAGGCCTGCAACCAGCATTTTCAATA 2212

Qy    267 IlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHis 285
          ||||||::::::||||||   :::|||      :::   |||      |||   :::
Db   2213 ATATTTTCAAGGATTATAGGGATCTTTACCCGAGATGAGGATCCTGAAACAAAACGACAG 2272

Qy    286 AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr 305
          ::::::   ::::::|||::::::|||::::::||||||:::|||   |||:::::::::
Db   2273 AATAGTAACATGTTTTCTGTATTGTTTCTAGTCCTTGGAATTATTTCTTTTATTACATTT 2332
```

```
Qy    306 PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHis 325
Db   2333 TTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTCACTAAGCGGCTTCGATAC 2392

Qy    326 LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSer 345
Db   2393 ATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTTGATGACCCTAAAAACACC 2452

Qy    346 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365
Db   2453 ACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAAGTTAAAGGGGCTATAGGT 2512

Qy    366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385
Db   2513 TCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGGACAGGCATTATTATATCC 2572

Qy    386 PheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405
Db   2573 TTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATTGTACCCATCATTGCAATA 2632

Qy    406 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425
Db   2633 GCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAGCTA 2692

Qy    426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445
Db   2693 GAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTCCGAACTGTTGTTTCTTTG 2752

Qy    446 ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn 465
Db   2753 ACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTGCAAGTACCATACAGAAAC 2812

Qy    466 ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr 485
Db   2813 TCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATCACCCAGGCAATGATGTAT 2872

Qy    486 PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr 505
Db   2873 TTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTGGCAAATGAGTTCATGAAC 2932

Qy    506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys 525
Db   2933 TTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAG 2992

Qy    526 ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAla 545
Db   2993 GTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCAGCAGCCCACGTCATCATG 3052

Qy    546 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 565
Db   3053 ATCATTGAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCACGGCCTCAAGCCAAATACG 3112

Qy    566 CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspVal 585
Db   3113 TTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTATCCCACTCGACCAGACATC 3172

Qy    586 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605
Db   3173 CCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAGACGCTGGCCCTCGTAGGT 3232

Qy    606 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProVal 625
Db   3233 AGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAGCGCTTCTATGACCCCTTG 3292

Qy    626 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArg 645
Db   3293 GCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTGAATGTCCAGTGGCTCCGA 3352

Qy    646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665
Db   3353 GCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGACTGCAGCATTGCCGAGAAC 3412

Qy    666 IleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn 685
Db   3413 ATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAGATTATGCAGGCAGCCAAG 3472

Qy    686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 705
Db   3473 GAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAATACAACACCAGAGTAGGA 3532

Qy    706 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725
Db   3533 GACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATTGCCATAGCTCGCGCTCTT 3592

Qy    726 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745
Db   3593 GTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAGT 3652

Qy    746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765
Db   3653 GAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATC 3712

Qy    766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785
Db   3713 GCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTGGTGTTTCAGAATGGCAAA 3772

Qy    786 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 805
Db   3773 GTCAAGGAGCATGGCACACATCAACAGCTGCTGGCCCAGAAAGGCATCTATTTTTCCATG 3832

Qy    806 ValAsnAlaGlnSer 810
Db   3833 GTCAGTGTCCAGGCT 3847
```

```
RESULT 7
AAD03505
ID   AAD03505 standard; cDNA; 4279 BP.
XX
AC   AAD03505;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.
XX
KW   Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KW   MDR1; drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag=  a
FT                   /product= "Dog P-glycoprotein (PGP) allelic variant
FT                   (Genotype B) protein"
FT   allele          replace (91, T)
FT                   /*tag=  b
FT   allele          replace (607, C)
FT                   /*tag=  c
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00309.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
```

```
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 9; Page 93-99; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) allelic variant
CC   (Genotype B) cDNA. The PGP enzyme functions as an efflux pump
CC   exporting small molecules across the cell membrane. This enzyme
CC   is a member of the ABC transporter family.
XX
SQ   Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;

Alignment Scores:
Pred. No.:              1.69e-226    Length:         4279
Score:                  2442.00      Matches:        469
Percent Similarity:     76.24%       Conservative:   160
Best Local Similarity:  56.85%       Mismatches:     178
Query Match:            59.87%       Indels:         18
DB:                     22           Gaps:           4

US-09-873-409-2 (1-812) x AAD03505 (1-4279)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db   1379 ATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTG 1438

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db   1439 GTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGC 1498

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db   1499 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTT 1558

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db   1559 ATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGT 1618

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db   1619 GGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTT 1678

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db   1679 CTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCC 1738

Qy    122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db   1739 CTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCTCATCGTTTGTCTACAGTT 1798

Qy    142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db   1799 CGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 1858

Qy    162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLys 181
Db   1859 GATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTCACAATGCAG---ACAAGA 1915

Qy    182 LysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu 201
Db   1916 GGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAAAGTGAAAGTGATGCCTTG 1975

Qy    202 ProLeuHisSerValLysSerIleLysSerAspPheIle-------------------- 214
Db   1976 GAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAAAGAAGATCAACTCGCAGG 2032

Qy    215 -----------------------AspLysAlaGluGluSerThrGlnSerLysGluIle 226
Db   2033 AGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACAAAAGAGGACTTGAATGAG 2092

Qy    227 SerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPhe 246
Db   2093 AATGTACCTCCAGTTTCCTTCTGGAGGATTCTGAAGCTGAACTCAACTGAATGGCCTTAT 2152

Qy    247 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 266
Db   2153 TTTGTGGTTGGTATATTTGTGCTATTATAAACGGAGGCCTGCAACCAGCATTTTCAATA 2212

Qy    267 IlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHis 285
Db   2213 ATATTTTCAAGGATTATAGGGATCTTTACCCGAGATGAGGATCCTGAAACAAAACGACAG 2272

Qy    286 AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr 305
Db   2273 AATAGTAACATGTTTTCTGTATTGTTTCTAGTCCTTGGAATTATTTCTTTTATTACATTT 2332

Qy    306 PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHis 325
Db   2333 TTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTCACTAAGCGGCTTCGATAC 2392

Qy    326 LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSer 345
Db   2393 ATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTTGATGACCCTAAAAACACC 2452

Qy    346 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365
Db   2453 ACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAAGTTAAAGGGGCTATAGGT 2512

Qy    366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385
Db   2513 TCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGGACAGGCATTATTATATCC 2572

Qy    386 PheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405
Db   2573 TTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATTGTACCCATCATTGCAATA 2632

Qy    406 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425
Db   2633 GCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAGCTA 2692

Qy    426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445
Db   2693 GAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTCCGAACTGTTGTTTCTTTG 2752

Qy    446 ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn 465
Db   2753 ACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTGCAAGTACCATACAGAAAC 2812

Qy    466 ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr 485
Db   2813 TCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATCACCCAGGCAATGATGTAT 2872

Qy    486 PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr 505
Db   2873 TTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTGGCAAATGAGTTCATGAAC 2932

Qy    506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys 525
Db   2933 TTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAG 2992

Qy    526 ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAla 545
Db   2993 GTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCAGCAGCCCACGTCATCATG 3052
```

```
Qy     546 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 565
           ::::::||||||   |||  ||||||||   |||        |||   ||||||:::|||
Db    3053 ATCATTGAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCACGGCCTCAAGCCAAATACG 3112

Qy     566 CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspVal 585
              |||||||||:::   |||   ||||||   |||   ||||||   |||||||||:::
Db    3113 TTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTATCCCACTCGACCAGACATC 3172

Qy     586 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605
              :::|||:::||||||||||   :::::::::|||:::|||:::|||   ||||||
Db    3173 CCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAGACGCTGGCCCTCGTAGGT 3232

Qy     606 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProVal 625
           |||||||||||||||||||||||||||   ||||||||||||:::|||   |||||||||:::
Db    3233 AGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAGCGCTTCTATGACCCCTTG 3292

Qy     626 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArg 645
              |||   ||||||   ||||||   :::   |||   ||||||||||||||||||
Db    3293 GCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTGAATGTCCAGTGGCTCCGA 3352

Qy     646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665
           :::   :::   ||||||   ||||||||||:::||||||:::||||||||||||||||
Db    3353 GCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGACTGCAGCATTGCCGAGAAC 3412

Qy     666 IleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn 685
           ||||||||||||||||||||||||||||         :::||||||   :::||||||
Db    3413 ATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAGATTATGCAGGCAGCCAAG 3472

Qy     686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 705
              ||||||||||||   |||||||||   |||||||||||||||||||||:::||||||
Db    3473 GAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAATACAACACCAGAGTAGGA 3532

Qy     706 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725
              ||||||   |||||||||||||||||||||||:::||||||||||||||||||
Db    3533 GACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATTGCCATAGCTCGCGCTCTT 3592

Qy     726 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745
           :::::::::|||   ||||||||||||||||||||||||||||||||||   :::|||
Db    3593 GTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAGT 3652

Qy     746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765
           ||||||||||||||   |||||||||||||||||   |||||||||||:::|||:::
Db    3653 GAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATC 3712

Qy     766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785
              |||||||||||   |||||||||||||||||||||||||||      ||||||||
Db    3713 GCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTGGTGTTTCAGAATGGCAAA 3772

Qy     786 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 805
           :::||||||   ||||||||||||:::||||||         :::   |||||||||   :::
Db    3773 GTCAAGGAGCATGGCACACATCAACAGCTGCTGGCCCAGAAAGGCATCTATTTTTCCATG 3832

Qy     806 ValAsnAlaGlnSer 810
           |||:::   |||:::
Db    3833 GTCAGTGTCCAGGCT 3847
```

RESULT 8

```
AAF86127
ID   AAF86127 standard; cDNA; 4186 BP.
XX
AC   AAF86127;
XX
DT   25-JUN-2001  (first entry)
XX
DE   Cynomologous monkey P-glycoprotein cDNA.
XX
KW   Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   efflux pump; ss.
XX
OS   Macaca fascicularis.
XX
FH   Key             Location/Qualifiers
FT   CDS             100..3942
FT                   /*tag=  a
FT                   /product= "PGP"
FT                   /note= "P-glycoprotein"
XX
PN   WO200123565-A1.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26592.
XX
PR   28-SEP-1999;   99US-0156921.
PR   12-OCT-1999;   99US-0158818.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-Crespi DT,  Crespi CL;
XX
DR   WPI; 2001-316136/33.
DR   P-PSDB; AAB81064.
XX
PT   Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein
PT   (PGP) and homologous PGP polypeptides are useful for predicting
PT   bioavailability of compound and increasing PGP transporter activity in
PT   cell  -
XX
PS   Example 1; Page 51-57; 84pp; English.
XX
CC   This invention relates to a polynucleotide sequence encoding a
CC   cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the
CC   PGP protein. PGP, also known as multidrug transporter, MDR1 is a member
CC   of the ABC transporter superfamily. The enzyme serves as an efflux pump
CC   exporting small molecules across the cell membrane. The invention
CC   includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence
CC   and protein, and also that of an allelic variant. The PGP polynucleotide
CC   sequence is useful for increasing PGP transporter activity in a cell.
CC   Antisense sequences of the cDNA are useful for inhibiting PGP transport
CC   activity in a mammalian cell. They may also be used for increasing the
CC   bioavailability of a drug. The present sequence represents cDNA encoding
CC   cynomologous monkey P-glycoprotein.
XX
SQ   Sequence 4186 BP; 1226 A; 799 C; 1039 G; 1122 T; 0 other;

Alignment Scores:
Pred. No.:              2.04e-226      Length:        4186
Score:                  2441.00        Matches:       465
Percent Similarity:     76.73%         Conservative:  168
Best Local Similarity:  56.36%         Mismatches:    174
Query Match:            59.84%         Indels:        18
DB:                     22             Gaps:          5

US-09-873-409-2 (1-812) x AAF86127 (1-4186)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
           ||||||   |||||||||   :::|||||||||   |||:::   |||||||||
Db    1456 GTTGATGGACAGGATATTAGGACCATAAACGTAAGGTTTCTACGGGAAATCATCGGTGTG 1515

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           ||||||||||||||||||||||||   |||||||||:::   ||||||:::|||||||||
Db    1516 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGTCGT 1575

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
           :::|||||||||   :::|||:::|||:::|||   :::||||||||||||||||||
Db    1576 GAAGATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1635

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           ||||||:::   |||   ||||||:::||||||||||||||||:::||||||||:::|||
Db    1636 ATCATGAAACTGCCTCAGAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGCTGAGT 1695

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1696 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 1755
```

```
Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db   1756 CTGCTGGACGAGGCCACGTCAGCCTTGGACACAGAAAGTGAAGCAGTGGTTCAGGTGGCT 1815

Qy    122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db   1816 CTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGTTTGTCTACGGTT 1875

Qy    142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db   1876 CGTAATGCCGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 1935

Qy    162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle--- 180
Db   1936 GATGAGCTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATGCAGACAGCAGGA 1995

Qy    181 -----------------LysLysAlaAspGlu---------GlnMetGluSerMetThr 191
Db   1996 AATGAAATTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATTGATACCTTGGAA 2055

Qy    192 TyrSerThrGluArgLysThrAsnSerLeu---------------ProLeuHisSerVal 206
Db   2056 ATGTCTTCACATGATTCAGGATCCAGTCTAATAAGAAAAAGATCCACTCGTAGGAGTGTC 2115

Qy    207 LysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIle 226
Db   2116 CGTGGATCACAAGGCCAA------GACAGAAAGCTTAGTACCAAAGAGGCTCTGGATGAA 2169

Qy    227 SerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPhe 246
Db   2170 AGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAGTGGCCTTAT 2229

Qy    247 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 266
Db   2230 TTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGTCTGCAACCAGCATTTGCAGTA 2289

Qy    267 IlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHis 285
Db   2290 ATATTTTCAAAGATTATAGGGATTTTTACAAGAAATGATGATGCCGAAACAAAACGACAG 2349

Qy    286 AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr 305
Db   2350 AATAGTAACTTGTTTTCACTATTGTTTCTAGTCCTTGGAATTGTTTCTTTTATTACATTT 2409

Qy    306 PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHis 325
Db   2410 TTCCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTCCGATAC 2469

Qy    326 LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSer 345
Db   2470 ATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGCTGGTTTGATGACCCTAAAAACACC 2529

Qy    346 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365
Db   2530 ACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGT 2589

Qy    366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385
Db   2590 TCCAGGCTTGCTATAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATTATATCC 2649

Qy    386 PheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405
Db   2650 TTAATCTATGGTTGGCAACTGACACTGTTACTCTTAGCAATTGTACCCATCATTGCAATA 2709

Qy    406 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425
Db   2710 GCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAACTA 2769

Qy    426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445
Db   2770 GAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACTGTTGTTTCTTTG 2829

Qy    446 ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn 465
Db   2830 ACTCAGGAGCAGAAGTTTGAACATATGTATGATCAGAGTTTGCAGGTACCATACAGAAAC 2889

Qy    466 ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr 485
Db   2890 TCTTTGAGGAAAGCACACATCTTTGGAATCACGTTTTCCTTCACGCAGGCAATGATGTAT 2949

Qy    486 PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr 505
Db   2950 TTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAGTCTCATGAGC 3009

Qy    506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys 525
Db   3010 TTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAA 3069

Qy    526 ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAla 545
Db   3070 GTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCAGCAGCCCACATCATCATG 3129

Qy    546 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 565
Db   3130 ATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACAGAAGGCCTAAAGCCGAACACA 3189

Qy    566 CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspVal 585
Db   3190 TTGGAAGGAAATGTCACATTTAATGAAGTTGTATTCAACTATCCCACCCGACTGGACATC 3249

Qy    586 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605
Db   3250 CCAGTGCTTCAGGGGCTGAGCCTGGAAGTGAAGAAGGGCCAGACGCTGGCCCTGGTGGGC 3309

Qy    606 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProVal 625
Db   3310 AGCAGTGGCTGTGGGAAGAGCACGGTGGTCCAGCTCCTGGAGCGGTTCTATGACCCCTTG 3369

Qy    626 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArg 645
Db   3370 GCGGGGAAAGTGCTGCTTGACGGCAAAGAAATAAAGCAACTGAATGTTCAGTGGCTCCGA 3429

Qy    646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665
Db   3430 GCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTAGTGAGAAC 3489

Qy    666 IleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn 685
Db   3490 ATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCAGCCAAG 3549

Qy    686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 705
Db   3550 GAGGCCAATATACACGCCTTCATCGAGTCACTGCCTAATAAATATAGCACCAGAGTAGGA 3609

Qy    706 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725
Db   3610 GACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTT 3669

Qy    726 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745
Db   3670 GTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACATCAGCTCTGGATACAGAAAGT 3729

Qy    746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765
Db   3730 GAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGTACCTGCATTGTGATT 3789

Qy    766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785
Db   3790 GCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGCAGA 3849

Qy    786 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 805
Db   3850 GTCAAGGAGCACGGCACACATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATG 3909

Qy    806 ValAsnAlaGlnSer 810
```

```
Db   3910 GTCAGTGTCCAGGCT 3924

RESULT 9
AAF86128
ID   AAF86128 standard; cDNA; 4195 BP.
XX
AC   AAF86128;
XX
DT   25-JUN-2001  (first entry)
XX
DE   Cynomologous monkey P-glycoprotein cDNA variant 1.
XX
KW   Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   efflux pump; ss.
XX
OS   Macaca fascicularis.
XX
FH   Key             Location/Qualifiers
FT   CDS             100..3951
FT                   /*tag= a
FT                   /product= "PGP"
FT                   /note= "P-glycoprotein"
FT   misc_feature    376..384
FT                   /*tag= b
FT                   /note= "Insertion of 9 nucleotides relative to PGP
FT                          allelic variant AAF86127"
XX
PN   WO200123565-A1.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26592.
XX
PR   28-SEP-1999;   99US-0156921.
PR   12-OCT-1999;   99US-0158818.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-Crespi DT,  Crespi CL;
XX
DR   WPI; 2001-316136/33.
DR   P-PSDB; AAB81065.
XX
PT   Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein
PT   (PGP) and homologous PGP polypeptides are useful for predicting
PT   bioavailability of compound and increasing PGP transporter activity in
PT   cell  -
XX
PS   Example 1; Page 59-65; 84pp; English.
XX
CC   This invention relates to a polynucleotide sequence encoding a
CC   cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the
CC   PGP protein. PGP, also known as multidrug transporter, MDR1 is a member
CC   of the ABC transporter superfamily. The enzyme serves as an efflux pump
CC   exporting small molecules across the cell membrane. The invention
CC   includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence
CC   and protein, and also that of an allelic variant. The PGP polynucleotide
CC   sequence is useful for increasing PGP transporter activity in a cell.
CC   Antisense sequences of the cDNA are useful for inhibiting PGP transport
CC   activity in a mammalian cell. They may also be used for increasing the
CC   bioavailability of a drug. The present sequence represents cDNA encoding
CC   cynomologous monkey P-glycoprotein. This sequence contains a 9 nucleotide
CC   insert compared to the PGP allelic variant given in AAF86127.
XX
SQ   Sequence 4195 BP; 1230 A; 801 C; 1039 G; 1125 T; 0 other;

Alignment Scores:
Pred. No.:               2.05e-226     Length:        4195
Score:                   2441.00       Matches:       465
Percent Similarity:      76.73%        Conservative:  168
Best Local Similarity:   56.36%        Mismatches:    174
Query Match:             59.84%        Indels:        18
DB:                      22            Gaps:          5

US-09-873-409-2 (1-812) x AAF86128 (1-4195)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
          |||||       |||||||||     :::|||||||||     |||:::   |||||||||
Db   1465 GTTGATGGACAGGATATTAGGACCATAAACGTAAGGTTTCTACGGGAAATCATCGGTGTG 1524

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
          ||||||||||||||||||||||||   |||||||||:::     ||||||:::|||||||||
Db   1525 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGTCGT 1584

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
          :::|||||||||   :::|||:::|||:::|||   :::|||||||||||||||||||||
Db   1585 GAAGATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1644

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
          ||||||:::   |||   ||||||:::|||||||||||||||:::|||||||||:::|||
Db   1645 ATCATGAAACTGCCTCAGAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGCTGAGT 1704

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1705 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 1764

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
          :::||||||||||||||||||||||||||||:::||||||::::::   ||||||   |||
Db   1765 CTGCTGGACGAGGCCACGTCAGCCTTGGACACAGAAAGTGAAGCAGTGGTTCAGGTGGCT 1824

Qy    122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
          |||:::||||||   ||||||||||||||||||||:::||||||||||||||||||:::
Db   1825 CTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGTTTGTCTACGGTT 1884

Qy    142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
          |||:::||||||:::|||             ||||||::::::   ||||||||   |||
Db   1885 CGTAATGCCGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 1944

Qy    162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle--- 180
             |||||||||   ::::::|||:::|||:::   ||||||      |||
Db   1945 GATGAGCTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATGCAGACAGCAGGA 2004

Qy    181 -----------------LysLysAlaAspGlu---------GlnMetGluSerMetThr 191
                           |||||||||         :::::::::::::::
Db   2005 AATGAAATTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATTGATACCTTGGAA 2064

Qy    192 TyrSerThrGluArgLysThrAsnSerLeu---------------ProLeuHisSerVal 206
             |||:::        :::||||||                        ||||||
Db   2065 ATGTCTTCACATGATTCAGGATCCAGTCTAATAAGAAAAAGATCCACTCGTAGGAGTGTC 2124

Qy    207 LysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIle 226
          :::      :::         |||:::         :::   ::::::   :::
Db   2125 CGTGGATCACAAGGCCAA------GACAGAAAGCTTAGTACCAAAGAGGCTCTGGATGAA 2178

Qy    227 SerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPhe 246
          |||:::|||   ||||||      :::|||:::|||||||||      |||||||||:::
Db   2179 AGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAGTGGCCTTAT 2238

Qy    247 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 266
             |||:::|||         :::::::::||||||   :::   |||   |||::::::
Db   2239 TTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGTCTGCAACCAGCATTTGCAGTA 2298

Qy    267 IlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHis 285
          ||||||:::||||||||   :::|||      |||:::|||      |||      :::
Db   2299 ATATTTTCAAAGATTATAGGGATTTTTACAAGAAATGATGATGCCGAAACAAAACGACAG 2358

Qy    286 AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr 305
          ::::::   ::::::|||::::::|||::::::||||||::::::   |||:::::::::
Db   2359 AATAGTAACTTGTTTTCACTATTGTTTCTAGTCCTTGGAATTGTTTCTTTTATTACATTT 2418

Qy    306 PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHis 325
          |||:::||||||      :::|||:::||||||||||||||||||   |||||||||:::
Db   2419 TTCCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTCCGATAC 2478

Qy    326 LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSer 345
          :::   |||::::::||||||   ||||||::::::||||||||:::   :::|||:::
```

```
Db   2479 ATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGCTGGTTTGATGACCCTAAAAACACC 2538

Qy    346 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365
          ||||||   |||||||||  ||||||  |||  ||||||::::::||||||   |||
Db   2539 ACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGT 2598

Qy    366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385
          ||||||:::  ::::::|||||||||      |||:::|||      :::|||||||||
Db   2599 TCCAGGCTTGCTATAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATTATATCC 2658

Qy    386 PheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405
             ||||||||||||:::::::|||   |||:::|||:::|||   |||::::::|||:::
Db   2659 TTAATCTATGGTTGGCAACTGACACTGTTACTCTTAGCAATTGTACCCATCATTGCAATA 2718

Qy    406 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425
             |||::::::|||         ::::::|||   |||   |||||||||:::||||||
Db   2719 GCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAACTA 2778

Qy    426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445
          :::   ||||||||||||||||||||||||:::||||||   ||||||:::|||||||||
Db   2779 GAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACTGTTGTTTCTTTG 2838

Qy    446 ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn 465
          |||:::|||:::   ||||||   ||||||::::::   ||||||      :::||||||
Db   2839 ACTCAGGAGCAGAAGTTTGAACATATGTATGATCAGAGTTTGCAGGTACCATACAGAAAC 2898

Qy    466 ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr 485
          :::   :::||||||   |||   |||      ::::::|||:::   |||   :::|||
Db   2899 TCTTTGAGGAAAGCACACATCTTTGGAATCACGTTTTCCTTCACGCAGGCAATGATGTAT 2958

Qy    486 PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr 505
          |||:::||||||      |||||||||||||||||||||:::            |||:::
Db   2959 TTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAGTCTCATGAGC 3018

Qy    506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys 525
             |||   :::   :::||||||:::|||:::   :::||||||||||||:::|||:::
Db   3019 TTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAA 3078

Qy    526 ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAla 545
                      ||||||:::|||:::|||||||||      |||||||||:::
Db   3079 GTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCAGCAGCCCACATCATCATG 3138

Qy    546 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 565
          ::::::||||||   |||   |||||||||   |||   ||||||:::|||
Db   3139 ATCATTGAAAAAAACCCCTTTGATTGACAGCTACAGCACAGAAGGCCTAAAGCCGAACACA 3198

Qy    566 CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspVal 585
             |||||||||:::   |||   ||||||   |||   ||||||   |||   |||:::
Db   3199 TTGGAAGGAAATGTCACATTTAATGAAGTTGTATTCAACTATCCCACCCGACTGGACATC 3258

Qy    586 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605
             :::|||:::||||||||||||   :::::::::|||:::|||:::|||   ||||||
Db   3259 CCAGTGCTTCAGGGGCTGAGCCTGGAAGTGAAGAAGGGCCAGACGCTGGCCCTGGTGGGC 3318

Qy    606 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProVal 625
          ||||||||||||||||||||||||   ||||||||||||:::|||   |||||||||:::
Db   3319 AGCAGTGGCTGTGGGAAGAGCACGGTGGTCCAGCTCCTGGAGCGGTTCTATGACCCCTTG 3378

Qy    626 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArg 645
             |||:::||||||   ||||||      :::   |||:::|||||||||||||||||||
Db   3379 GCGGGGAAAGTGCTGCTTGACGGCAAAGAAATAAAGCAACTGAATGTTCAGTGGCTCCGA 3438

Qy    646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665
          :::   :::   ||||||   |||||||||:::||||||:::|||||||||:::||||||
Db   3439 GCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTAGTGAGAAC 3498

Qy    666 IleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn 685
          ||||||||||||||||||||||||||||||      :::||||||      ||||||
Db   3499 ATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCAGCCAAG 3558

Qy    686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 705
             ||||||||||||:::|||||||||   ||||||   ||||||:::|||:::||||||
Db   3559 GAGGCCAATATACACGCCTTCATCGAGTCACTGCCTAATAAATATAGCACCAGAGTAGGA 3618

Qy    706 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725
             ||||||   |||||||||||||||||||||||||:::||||||||||||||||||
Db   3619 GACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTT 3678

Qy    726 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745
          :::::::::|||   ||||||||||||||||||||||||||||||||||||   :::|||
Db   3679 GTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACATCAGCTCTGGATACAGAAAGT 3738

Qy    746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765
          ||||||||||||||   ||||||||||||||||||   ||||||||||||:::|||:::
Db   3739 GAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGTACCTGCATTGTGATT 3798

Qy    766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785
             ||||||||||||   ||||||||||||||||||||||||||      ||||||:::
Db   3799 GCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGCAGA 3858

Qy    786 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 805
          :::||||||   |||||||||||||:::||||||      :::   |||||||||   :::
Db   3859 GTCAAGGAGCACGGCACACATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATG 3918

Qy    806 ValAsnAlaGlnSer 810
          |||:::   |||:::
Db   3919 GTCAGTGTCCAGGCT 3933
```

RESULT 10
AAD03506

```
ID   AAD03506 standard; cDNA; 4279 BP.
XX
AC   AAD03506;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) allelic variant (Genotype D) cDNA.
XX
KW   Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KW   MDR1; drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag= a
FT                   /product= "Dog P-glycoprotein (PGP) allelic variant
FT                   (Genotype D) protein"
FT   allele          replace (91, T)
FT                   /*tag= b
FT   allele          replace (607, C)
FT                   /*tag= c
FT   allele          replace (1001, T)
FT                   /*tag= c
FT   allele          replace (3458, A)
FT                   /*tag= c
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00310.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
```

```
PS   Claim 9; Page 102-107; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) allelic variant
CC   (Genotype D) cDNA. The PGP enzyme functions as an efflux pump
CC   exporting small molecules across the cell membrane. This enzyme
CC   is a member of the ABC transporter family.
XX
SQ   Sequence 4279 BP; 1296 A; 833 C; 1009 G; 1141 T; 0 other;

Alignment Scores:
Pred. No.:              2.11e-226     Length:         4279
Score:                  2441.00       Matches:        469
Percent Similarity:     76.24%        Conservative:   160
Best Local Similarity:  56.85%        Mismatches:     178
Query Match:            59.84%        Indels:         18
DB:                     22            Gaps:           4

US-09-873-409-2 (1-812) x AAD03506 (1-4279)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db   1379 ATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTG 1438

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db   1439 GTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGC 1498

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db   1499 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTT 1558

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db   1559 ATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGT 1618

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db   1619 GGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTT 1678

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db   1679 CTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCC 1738

Qy    122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db   1739 CTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCTCATCGTTTGTCTACAGTT 1798

Qy    142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db   1799 CGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 1858

Qy    162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLys 181
Db   1859 GATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTCACAATGCAG---ACAAGA 1915

Qy    182 LysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu 201
Db   1916 GGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAAAGTGAAAGTGATGCCTTG 1975

Qy    202 ProLeuHisSerValLysSerIleLysSerAspPheIle--------------------- 214
Db   1976 GAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAAAGAAGATCAACTCGCAGG 2032

Qy    215 -----------------------AspLysAlaGluGluSerThrGlnSerLysGluIle 226
Db   2033 AGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACAAAAGAGGACTTGAATGAG 2092

Qy    227 SerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPhe 246
Db   2093 AATGTACCTCCAGTTTCCTTCTGGAGGATTCTGAAGCTGAACTCAACTGAATGGCCTTAT 2152

Qy    247 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 266
Db   2153 TTTGTGGTTGGTATATTTGTGCTATTATAAACGGAGGCCTGCAACCAGCATTTTCAATA 2212

Qy    267 IlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHis 285
Db   2213 ATATTTTCAAGGATTATAGGGATCTTTACCCGAGATGAGGATCCTGAAACAAAACGACAG 2272

Qy    286 AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr 305
Db   2273 AATAGTAACATGTTTTCTGTATTGTTTCTAGTCCTTGGAATTATTTCTTTTATTACATTT 2332

Qy    306 PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHis 325
Db   2333 TTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTCACTAAGCGGCTTCGATAC 2392

Qy    326 LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSer 345
Db   2393 ATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTTGATGACCCTAAAAACACC 2452

Qy    346 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365
Db   2453 ACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAAGTTAAAGGGGCTATAGGT 2512

Qy    366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385
Db   2513 TCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGGACAGGCATTATTATATCC 2572

Qy    386 PheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405
Db   2573 TTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATTGTACCCATCATTGCAATA 2632

Qy    406 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425
Db   2633 GCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAGCTA 2692

Qy    426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445
Db   2693 GAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTCCGAACTGTTGTTTCTTTG 2752

Qy    446 ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn 465
Db   2753 ACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTGCAAGTACCATACAGAAAC 2812

Qy    466 ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr 485
Db   2813 TCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATCACCCAGGCAATGATGTAT 2872

Qy    486 PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr 505
Db   2873 TTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTGGCAAATGAGTTCATGAAC 2932

Qy    506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys 525
Db   2933 TTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAG 2992

Qy    526 ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAla 545
Db   2993 GTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCAGCAGCCCACGTCATCATG 3052

Qy    546 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 565
```

```
Db    3053 ATCATTGAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCACGGCCTCAAGCCAAATACG 3112

Qy     566 CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspVal 585
                 |||||||||:::   |||  ||||||  |||  ||||||   |||||||||:::
Db    3113 TTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTATCCCACTCGACCAGACATC 3172

Qy     586 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605
              :::|||:::||||||||||||   :::::::::|||:::|||:::|||   ||||||
Db    3173 CCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAGACGCTGGCCCTCGTAGGT 3232

Qy     606 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProVal 625
           ||||||||||||||||||||||||   ||||||||||||:::|||   |||||||||:::
Db    3233 AGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAGCGCTTCTATGACCCCTTG 3292

Qy     626 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArg 645
              |||  ||||||  ||||||  :::   |||   ||||||||||||||||||||||
Db    3293 GCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTGAATGTCCAGTGGCTCCGA 3352

Qy     646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665
           :::   :::   ||||||   ||||||||||:::||||||:::||||||||||||||||||
Db    3353 GCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGACTGCAGCATTGCCGAGAAC 3412

Qy     666 IleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn 685
           |||||||||||||||||||||||||||||||      :::||||||   :::||||||
Db    3413 ATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAGATTGTGCAGGCAGCCAAG 3472

Qy     686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 705
              ||||||||||||   |||||||||   |||||||||||||||||||:::||||||
Db    3473 GAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAATACAACACCAGAGTAGGA 3532

Qy     706 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725
              ||||||   ||||||||||||||||||||||||||:::|||||||||||||||||||
Db    3533 GACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATTGCCATAGCTCGCGCTCTT 3592

Qy     726 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745
           :::::::::|||   |||||||||||||||||||||||||||||||||   :::|||
Db    3593 GTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAGT 3652

Qy     746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765
           ||||||||||||||||   |||||||||||||||||||   ||||||||||||:::|||:::
Db    3653 GAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATC 3712

Qy     766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785
              ||||||||||||   ||||||||||||||||||||||||||   |||||||||
Db    3713 GCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTGGTGTTTCAGAATGGCAAA 3772

Qy     786 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 805
           :::||||||   ||||||||||||:::||||||      :::   |||||||||   :::
Db    3773 GTCAAGGAGCATGGCACACATCAACAGCTGCTGGCCCAGAAAGGCATCTATTTTTCCATG 3832

Qy     806 ValAsnAlaGlnSer 810
           |||:::   |||:::
Db    3833 GTCAGTGTCCAGGCT 3847

RESULT 11
AAZ52047
ID   AAZ52047 standard; cDNA; 4369 BP.
XX
AC   AAZ52047;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Rat multidrug resistance protein 1a cDNA.
XX
KW   Multidrug resistance protein 1a; mdr1a; multi-specific drug transporter;
KW   drug formulation; formulation excipient; compound design; inflammation;
KW   pharmacokinetic; oral absorption; cancer; cardiovascular disease;
KW   central nervous system disorder; auto-immune disease; kidney disease; ss.
XX
OS   Rattus rattus.
XX
FH   Key             Location/Qualifiers
FT   CDS             352..4170
FT                   /*tag= a
FT                   /product= "Rat multidrug resistance protein"
XX
PN   WO200015650-A1.
XX
PD   23-MAR-2000.
XX
PF   10-SEP-1999;   99WO-US20770.
XX
PR   17-SEP-1998;   98US-0156800.
PR   09-DEC-1998;   98US-0208809.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI   Brun KA,  Chenery RJ,  Ellens H,  Feild JA,  Yue L;
XX
DR   WPI; 2000-271372/23.
DR   P-PSDB; AAY70596.
XX
PT   Isolated rat mdr1a polynucleotides and polypeptides, useful in assays
PT   to provide information on drug formulation, selection of formulation
PT   excipients and compound design -
XX
PS   Claim 2; Page 23-24; 33pp; English.
XX
CC   The present cDNA sequence encodes rat multidrug resistance
CC   protein 1a (mdr1a). This protein is a member of the multi-specific drug
CC   transporters family. Mdr1a is used in assays to provide information on
CC   drug formulation, selection of formulation excipients and compound
CC   design. They are used in cell based, membrane based, binding or other
CC   assays to provide information that may enhance drug formulation. This
CC   invention further relates to the generation of in vivo and in vitro
CC   comparison data to predict oral absorption and pharmacokinetics. This
CC   enables the selection of drugs with optimal pharmacokinetics, i.e. good
CC   oral bioavailability, brain penetration, plasma half life, and minimum
CC   drug interaction. Transgenic and knock-out animals created using rat
CC   mdr1a provides an insight into treating and preventing human diseases
CC   including cancer, inflammation, cardiovascular disease, central nervous
CC   system disorders, auto-immune and kidney disease.
XX
SQ   Sequence 4369 BP; 1241 A; 973 C; 1096 G; 1059 T; 0 other;

Alignment Scores:
Pred. No.:              2.18e-226      Length:         4369
Score:                  2441.00        Matches:        470
Percent Similarity:     77.04%         Conservative:   164
Best Local Similarity:  57.11%         Mismatches:     175
Query Match:            59.84%         Indels:         14
DB:                     21             Gaps:           5

US-09-873-409-2 (1-812) x AAZ52047 (1-4369)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
           :::|||      ||||||||||   :::||||||||||:::   |||:::   |||||||||
Db    1684 ATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTG 1743

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           ||||||||||||||||||||||||   ||||||||||:::   ||||||:::|||||||||
Db    1744 GTGAGTCAGGAACCCGTGCTGTTTGCCACCACAATTGCCGAAAACATTCGCTATGGCCGA 1803

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
           ::::::||||||   :::|||:::|||:::|||   :::||||||||||||||||||||
Db    1804 GAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAATGCCTATGATTTC 1863

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           |||||||:::   |||:::|||||||:::||||||||||||||||:::||||||||||:::|||
Db    1864 ATCATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGT 1923

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1924 GGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAACCCCAAGATCCTT 1983

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
```

```
Db   1984 TTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTGGTTCAGGCCGCT 2043

Qy    122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db   2044 CTGGATAAGGCTAGAGAAGGCCGGACCACCATTGTGATAGCTCACCGCTTGTCTACAGTT 2103

Qy    142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db   2104 CGCAATGCTGACGTCATTGCTGGTTTGATGGTGGTGTCATTGTGGAGCAAGGAAATCAT 2163

Qy    162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178
Db   2164 GATGAGCTCATGAGAGAGAAAGGAATTTACTTCAAACTTGTCATGACTCAGACAGCAGGA 2223

Qy    179 ---AspIleLysLysAlaAspGluGlnMetGluSer-----------------MetThr 191
Db   2224 AATGAAATTGAATTAGGAAATGAAGCTTGTGAATCTAAAGACGGAATTGATAATGTGGAC 2283

Qy    192 TyrSerThrGluArgLysThrAsnSerLeu------ProLeuHisSerValLysSerIle 209
Db   2284 ATGTCTTCAAAAGATTCGGGATCCAGTCTAATAAGAAGAAGATCAACTCGCAAAAGCATC 2343

Qy    210 LysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLysGluIleSerLeu 228
Db   2344 CGTGGGCCACATGATCAAGACGGGGAACTTAGCACCAAAGAGGCTCTGGATGACGACGTA 2403

Qy    229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
Db   2404 CCTCCAGCTTCCTTTTGGCGGATCCTGAAGTTGAATTCAACTGAATGGCCTTATTTTGTG 2463

Qy    249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db   2464 GTTGGTGTATTTTGTGCCATAATAAATGGAGGCTTGCAACCAGCATTCTCCATAATATTT 2523

Qy    269 AlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeu---LysHisAspAla 287
Db   2524 TCAAAGGTTGTAGGGGTTTTTACAAAAAATGACACCCCTGAAATCCAGCGGCAGAACAGC 2583

Qy    288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db   2584 AACTTGTTTTCTTTATTGTTCCTGATCCTTGGGATCATCTCTTTCATTACGTTTTTCCTT 2643

Qy    308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db   2644 CAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGACTCCGATACATGGTC 2703

Qy    328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
Db   2704 TTCAAATCCATGCTGAGACAGGACATAAGCTGGTTTGATGACCCTAAAAACACCACAGGA 2763

Qy    348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db   2764 GCGCTGACCACCAGGCTTGCCAATGACGCTGCTCAAGTGAAAGGGGCTACAGGGTCTAGG 2823

Qy    368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db   2824 CTTGCTGTTATTACCCAGAACATAGCAAATCTTGGGACAGGCATCATCATATCCCTGATC 2883

Qy    388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db   2884 TACGGCTGGCAATTGACACTTTTACTCCTAGCAATTGTTCCCATCATTGCTATAGCAGGA 2943

Qy    408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db   2944 GTGGTTGAAATGAAAATGTTGTCTGGACAAGCGCTGAAAGATAAGAAGGAACTAGAAGGT 3003

Qy    428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db   3004 TCTGGGAAGATCGCTACAGAAGCAATTGAAAACTTTCGCACTGTCGTCTCTTTGACTCGG 3063

Qy    448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db   3064 GAGCAGAAGTTTGAAACTATGTATGCCCAGAGCTTGCAGATACCATACAGAAATGCTTTG 3123

Qy    468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db   3124 AAGAAAGCGCACGTCTTTGGGATCACTTTCTCCTTCACCCAGGCCATGATGTATTTCTCC 3183

Qy    488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db   3184 TATGCTGCTTGTTTCCGGTTTGATGCCTACTTGGTGGCACGAGAACTCATGACATTTGAA 3243

Qy    508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db   3244 AATGTTCTGTTAGTATTCTCAGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAGGTCAGT 3303

Qy    528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db   3304 TCATTCGCTCCTGACTACGCGAAAGCCAAAGTCTCGGCATCCCACATCATCAGGATCATT 3363

Qy    548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db   3364 GAGAAAATCCCTGAGATTGACAGCTACAGCACGGAGGGCTTGAAGCCTAATATGTTGGAA 3423

Qy    568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db   3424 GGAAATGTGAAATTTAATGGAGTCATGTTCAACTATCCCACCCGACCCAACATCCCAGTG 3483

Qy    588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db   3484 CTTCAGGGGCTGAGCCTAGAGGTGAAGAAAGGGCAGACGCTGGCCCTCGTGGGCAGCAGT 3543

Qy    608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db   3544 GGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTTGAGCGCTTCTATGACCCCATGGCCGGA 3603

Qy    628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
Db   3604 ACAGTGTTTCTAGATGGCAAAGAAATAAAGCAACTCAATGTCCAGTGGCTCCGCGCCCAC 3663

Qy    648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db   3664 CTGGGCATTGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATCGCCGAGAACATTGCC 3723

Qy    668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db   3724 TACGGAGACAACAGCCGTGTCGTGTCTCATAAGGAGATCGTGAAGGCAGCCAAGGAGGCC 3783

Qy    688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db   3784 AACATCCACCAGTTCATCGACTCACTGCCTGAGAAATACAACACCAGAGTGGGAGACAAA 3843

Qy    708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
Db   3844 GGGACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATCGCCATCGCGCGCGCCCTCGTCAGA 3903

Qy    728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db   3904 CAGCCTCACATCTTACTTCTGGATGAAGCGACATCAGCTCTGGATACGGAGAGTGAAAAG 3963

Qy    748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db   3964 GTCGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACCTGCATTGTGATCGCGCAC 4023

Qy    768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db   4024 CGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATTCAGAACGGCCAGGTCAAG 4083

Qy    788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db   4084 GAGCACGGCACCCACCAGCAGCTGCTGGCCCAGAAAGGCATCTATTTCTCGATGGTCAGT 4143

Qy    808 AlaGlnSer 810
Db   4144 GTGCAGGCT 4152
```

```
RESULT 12
AAZ52048
ID   AAZ52048 standard; cDNA; 4425 BP.
XX
AC   AAZ52048;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Rat multidrug resistance protein 1a cDNA derived from EST sequences.
XX
KW   Multidrug resistance protein 1a; mdr1a; multi-specific drug transporter;
KW   drug formulation; formulation excipient; compound design; inflammation;
KW   pharmacokinetic; oral absorption; cancer; cardiovascular disease;
KW   central nervous system disorder; auto-immune disease; kidney disease;
KW   EST; expressed sequence tag; ss.
XX
OS   Rattus rattus.
XX
FH   Key             Location/Qualifiers
FT   CDS             352..4170
FT                   /*tag=  a
FT                   /product= "Rat multidrug resistance protein"
XX
PN   WO200015650-A1.
XX
PD   23-MAR-2000.
XX
PF   10-SEP-1999;   99WO-US20770.
XX
PR   17-SEP-1998;   98US-0156800.
PR   09-DEC-1998;   98US-0208809.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI   Brun KA,  Chenery RJ,  Ellens H,  Feild JA,  Yue L;
XX
DR   WPI; 2000-271372/23.
DR   P-PSDB; AAY70597.
XX
PT   Isolated rat mdr1a polynucleotides and polypeptides, useful in assays
PT   to provide information on drug formulation, selection of formulation
PT   excipients and compound design -
XX
PS   Claim 11; Page 27-29; 33pp; English.
XX
CC   The present cDNA sequence encodes rat multidrug resistance protein 1a
CC   (mdr1a). This cDNA is derived from EST (expressed sequence tag)
CC   sequences. Mdr1a is used in assays to provide information on
CC   drug formulation, selection of formulation excipients and compound
CC   design. They are used in cell based, membrane based, binding or other
CC   assays to provide information that may enhance drug formulation. This
CC   invention further relates to the generation of in vivo and in vitro
CC   comparision data to predict oral absorption and pharmacokinetics. This
CC   enables the selection of drugs with optimal pharmacokinetics, i.e. good
CC   oral bioavailability, brain penetration, plasma half life, and minimum
CC   drug interaction. Transgenic and knock-out animals created using rat
CC   mdr1a provides an insight into treating and preventing human diseases
CC   including cancer, inflammation, cardiovascular disease, central nervous
CC   system disorders, auto-immune and kidney disease.
XX
SQ   Sequence 4425 BP; 1250 A; 992 C; 1113 G; 1067 T; 2 U; 1 other;

Alignment Scores:
Pred. No.:              2.22e-226     Length:        4425
Score:                  2441.00       Matches:       470
Percent Similarity:     77.04%        Conservative:  164
Best Local Similarity:  57.11%        Mismatches:    175
Query Match:            59.84%        Indels:        14
DB:                     21            Gaps:          5

US-09-873-409-2 (1-812) x AAZ52048 (1-4425)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
           :::|||      ||||||||   :::|||||||||:::   |||:::   |||||||||
Db    1684 ATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTG 1743

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           ||||||||||||||||||||||||   |||||||||:::   ||||||:::|||||||||
Db    1744 GTGAGTCAGGAACCCGTGCTGTTTGCCACCACAATTGCCGAAAACATTCGCTATGGCCGA 1803

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
           ::::::||||||   :::|||:::|||:::|||   :::||||||||||||||||||||||
Db    1804 GAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAATGCCTATGATTTC 1863

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           ||||||:::   |||:::||||||:::|||||||||||||||:::||||||||||:::|||
Db    1864 ATCATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGT 1923

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1924 GGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAACCCCAAGATCCTT 1983

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
           :::|||||||||||||||||||||||||||||||||:::||||||::::::   ||||||||||||
Db    1984 TTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTGGTTCAGGCCGCT 2043

Qy     122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
           |||:::||||||   :::|||||||||||||||||||:::|||||||||||||||||||:::
Db    2044 CTGGATAAGGCTAGAGAAGGCCGGACCACCATTGTGATAGCTCACCGCTTGTCTACAGTT 2103

Qy     142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
           |||:::||||||:::|||                  |||::::::   |||:::|||   |||
Db    2104 CGCAATGCTGAUGTCATTGCTGGTTTTGATGGTGGTGTCATTGTGGAGCAAGGAAATCAT 2163

Qy     162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178
              |||||||||   ::::::|||:::|||:::   |||||||||:::|||
Db    2164 GATGAGCTCATGAGAGAGAAAGGAATTTACTTCAAACTTGTCATGACTCAGACAGCAGGA 2223

Qy     179 ---AspIleLysLysAlaAspGluGlnMetGluSer------------------MetThr 191
              :::|||:::      :::|||      ||||||                   :::
Db    2224 AATGAAATTGAATTAGGAAATGAAGCTTGTGAATCTAAAGATGGAATTGATAATGTGGAC 2283

Qy     192 TyrSerThrGluArgLysThrAsnSerLeu------ProLeuHisSerValLysSerIle 209
              |||::::::      :::||||||             :::   |||||||||
Db    2284 ATGTCTTCAAAAGATTCRGGATCCAGTCTAATAAGAAGAAGATCAACTCGCAAAAGCATC 2343

Qy     210 LysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLysGluIleSerLeu 228
           :::                        |||   ||||||   ::::::   :::      :::
Db    2344 CGTGGGCCACATGATCAAGACGGGGAACTTAGCACCAAAGAGGCTCTGGATGACGACGTA 2403

Qy     229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
           |||      |||      :::|||||||||||||||      |||||||||:::   |||
Db    2404 CCTCCAGCTTCCTTTGGCGGATCCTGAAGTTGAATTCAACTGAATGGCCTTATTTTGTG 2463

Qy     249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
           :::|||         :::::::::||||||      :::      |||      |||||||||||||||
Db    2464 GTTGGTGTATTTTGTGCCATAATAAATGGAGGCTTGCAACCAGCATTCTCCATAATATTT 2523

Qy     269 AlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeu---LysHisAspAla 287
           :::|||::::::      :::|||      ||||||         :::   :::   ::::::
Db    2524 TCAAAGGTTGTAGGGGTTTTTACAAAAAATGACACCCCTGAAATCCAGCGGCAGAACAGC 2583

Qy     288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
              ::::::|||::::::|||:::|||||||||:::|||   |||:::::::::|||:::
Db    2584 AACTTGTTTTCTTTATTGTTCCTGATCCTTGGGATCATCTCTTTCATTACGTTTTTCCTT 2643

Qy     308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
           ||||||      :::|||:::||||||||||||||||||   |||||||||::::::
Db    2644 CAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGACTCCGATACATGGTC 2703

Qy     328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
           ||||||:::||||||   |||||||||:::|||||||||:::      :::|||:::||||||
Db    2704 TTCAAATCCATGCTGAGACAGGACATAAGCTGGTTTGATGACCCTAAAAACACCACAGGA 2763

Qy     348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
```

```
                 |||||||||   ||||||   |||   ||||||::::::||||||||||||||||
Db    2764 GCGCTGACCACCAGGCTTGCCAATGACGCTGCTCAAGTGAAAGGGGCTACAGGGTCTAGG 2823

Qy     368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
           :::   |||:::||||||||         |||:::|||     :::||||||||    |||
Db    2824 CTTGCTGTTATTACCCAGAACATAGCAAATCTTGGGACAGGCATCATCATATCCCTGATC 2883

Qy     388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
           |||||||||::::::|||      |||:::|||:::|||   |||::::::|||:::   |||
Db    2884 TACGGCTGGCAATTGACACTTTTACTCCTAGCAATTGTTCCCATCATTGCTATAGCAGGA 2943

Qy     408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
           ::::::|||        ::::::|||   |||   |||||||||:::||||||:::
Db    2944 GTGGTTGAAATGAAAATGTTGTCTGGACAAGCGCTGAAAGATAAGAAGGAACTAGAAGGT 3003

Qy     428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
           :::||||||||||||||||||||||:::||||||      ||||||:::|||||||||||||||
Db    3004 TCTGGGAAGATCGCTACAGAAGCAATTGAAAACTTTCGCACTGTCGTCTCTTTGACTCGG 3063

Qy     448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
           |||:::   ||||||   ||||||    :::   ||||||      :::||||||
Db    3064 GAGCAGAAGTTTGAAACTATGTATGCCCAGAGCTTGCAGATACCATACAGAAATGCTTTG 3123

Qy     468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
           |||||||||    :::   |||       ::::::|||:::    |||    :::||||||:::
Db    3124 AAGAAAGCGCACGTCTTTGGGATCACTTTCTCCTTCACCCAGGCCATGATGTATTTCTCC 3183

Qy     488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
           |||||||||   |||||||||   |||||||||:::              ||||||   |||
Db    3184 TATGCTGCTTGTTTCCGGTTTGATGCCTACTTGGTGGCACGAGAACTCATGACATTTGAA 3243

Qy     508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
              :::    :::||||||:::||||||    :::||||||||||||:::|||:::
Db    3244 AATGTTCTGTTAGTATTCTCAGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAGGTCAGT 3303

Qy     528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
                 ||||||:::|||:::|||||||||        |||:::|||:::        ::::::
Db    3304 TCATTCGCTCCTGACTACGCGAAAGCCAAAGTCTCGGCATCCCACATCATCAGGATCATT 3363

Qy     548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
           ||||||   |||   |||||||||   |||   ||||||   ||||||:::         |||
Db    3364 GAGAAAATCCCTGAGATTGACAGCTACAGCACGGAGGGCTTGAAGCCTAATATGTTGGAA 3423

Qy     568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
           ||||||::::::|||      |||   ||||||   ||||||::::::      :::
Db    3424 GGAAATGTGAAATTTAATGGAGTCATGTTCAACTATCCCACCCGACCCAACATCCCAGTG 3483

Qy     588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
           |||:::|||||||||||||   :::::::::|||:::|||:::|||   |||||||||||||
Db    3484 CTTCAGGGGCTGAGCCTAGAGGTGAAGAAAGGGCAGACGCTGGCCCTCGTGGGCAGCAGT 3543

Qy     608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
           ||||||||||||||||||   ||||||||||||:::|||   |||||||||:::   |||
Db    3544 GGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTTGAGCGCTTCTATGACCCCATGGCCGGA 3603

Qy     628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
              |||   ||||||    :::   |||:::||||||||||||||||||||||||:::
Db    3604 ACAGTGTTTCTAGATGGCAAAGAAATAAAGCAACTCAATGTCCAGTGGCTCCGCGCCCAC 3663

Qy     648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
           :::   ||||||   ||||||||||:::||||||:::||||||||||||||||||||||
Db    3664 CTGGGCATTGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATCGCCGAGAACATTGCC 3723

Qy     668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
           ||||||||||||||||||||||||        ||||||    :::||||||      |||
Db    3724 TACGGAGACAACAGCCGTGTCGTGTCTCATAAGGAGATCGTGAAGGCAGCCAAGGAGGCC 3783

Qy     688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
           |||||||||   ||||||:::   ||||||||||||||||||||||:::||||||    |||
Db    3784 AACATCCACCAGTTCATCGACTCACTGCCTGAGAAATACAACACCAGAGTGGGAGACAAA 3843

Qy     708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
           |||   |||||||||||||||||||||||||:::||||||||||||||||||::::::
Db    3844 GGGACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATCGCCATCGCGCGCGCCCTCGTCAGA 3903

Qy     728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
           :::|||   ||||||||||||||||||||||||||||||||||     :::|||||||||
Db    3904 CAGCCTCACATCTTACTTCTGGATGAAGCGACATCAGCTCTGGATACGGAGAGTGAAAAG 3963

Qy     748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
           |||||||||   ||||||||||||||||||   ||||||||||||:::|||:::   |||
Db    3964 GTCGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACCTGCATTGTGATCGCGCAC 4023

Qy     768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
           |||||||||   ||||||||||||||||||||||||||||:::      ||||||::::::|||
Db    4024 CGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATTCAGAACGGCCAGGTCAAG 4083

Qy     788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
           |||   ||||||||||||:::||||||         :::   |||||||||    :::|||:::
Db    4084 GAGCACGGCACCCACCAGCAGCTGCTGGCCCAGAAAGGCATCTATTTCTCGATGGTCAGT 4143

Qy     808 AlaGlnSer 810
              |||:::
Db    4144 GTGCAGGCT 4152
```

RESULT 13
AAD03489
ID   AAD03489 standard; cDNA; 4317 BP.
XX
AC   AAD03489;
XX
DT   13-JUN-2001 (first entry)
XX
DE   Dog P-glycoprotein (PGP) cDNA #2.
XX
KW   Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key   Location/Qualifiers
FT   CDS   70..3912
FT   /*tag= a
FT   /product= "Dog P-glycoprotein (PGP) #2"
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00304.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 1; Page 66-72; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP

```
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) cDNA. The
CC   PGP enzyme functions as an efflux pump exporting small molecules
CC   across the cell membrane. This enzyme is a member of the ABC
CC   transporter family.
XX
SQ   Sequence 4317 BP; 1293 A; 844 C; 1019 G; 1161 T; 0 other;

Alignment Scores:
Pred. No.:              4.18e-226      Length:         4317
Score:                  2438.00        Matches:        467
Percent Similarity:     76.24%         Conservative:   162
Best Local Similarity:  56.61%         Mismatches:     178
Query Match:            59.77%         Indels:         18
DB:                     22             Gaps:           4

US-09-873-409-2 (1-812) x AAD03489 (1-4317)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
           :::|||        |||||||||   :::|||||||||||||   |||:::    ||||||
Db   1429 ATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTG 1488

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
          ||||||||||||||||||||||||   ||||||||||:::   |||||:::||||||||
Db   1489 GTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGC 1548

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
          ::::::||||||   :::|||:::|||:::|||   :::||||||||||||||||||||||
Db   1549 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTT 1608

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
          ||||||:::   ||||||||||||:::||||||||||||||:::||||||:::::::|||
Db   1609 ATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAGAGAGGGGCCCGGCTGAGT 1668

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1669 GGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTT 1728

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
          :::||||||||||||||||||||||||||||:::||||||::::::   ||||||   |||
Db   1729 CTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCC 1788

Qy    122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
          |||:::||||||   ||||||||||||||||||||:::||||||||||||||||||:::
Db   1789 CTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCTCATCGTTTGTCTACAGTT 1848

Qy    142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
          |||:::||||||:::|||               ||||||::::::   ||||||||   |||
Db   1849 CGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 1908

Qy    162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLys 181
             |||||||||   ::::::|||:::|||:::   ||||||      |||         :::
Db   1909 GATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTCACAATGCAG---ACAAGA 1965

Qy    182 LysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu 201
             :::   ::::::|||:::   |||      ::::::   ::::::::::::|||
Db   1966 GGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAAAGTGAAAGTGATGCCTTG 2025

Qy    202 ProLeuHisSerValLysSerIleLysSerAspPheIle--------------------- 214
             :::   |||   |||         |||      |||
Db   2026 GAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAAAGAAGATCAACTCGCAGG 2082

Qy    215 ------------------------AspLysAlaGluGluSerThrGlnSerLysGluIle 226
                                  |||:::         :::   :::
Db   2083 AGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACAAAAGAGGACTTGAATGAG 2142

Qy    227 SerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPhe 246
          ::::::|||         ||||||   :::|||||||||||||||   |||||||||:::
Db   2143 AATGTTCCTTCAGTTTCCTTCTGGAGGATTCTGAAGCTGAACTCAACTGAATGGCCTTAT 2202

Qy    247 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 266
          |||:::|||         :::::::::||||||   :::   |||   |||||||||
Db   2203 TTTGTGGTTGGTATATTTTGTGCTATTATAAACGGAGGCCTGCAACCAGCATTTTCAATA 2262

Qy    267 IlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHis 285
          ||||||::::::||||||   :::|||      :::   |||      |||   :::
Db   2263 ATATTTTCAAGGATTATAGGGATCTTTACCCGAGATGAGGATCCTGAAACAAAACGACAG 2322

Qy    286 AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr 305
          ::::::   ::::::|||::::::|||::::::||||||:::|||   |||:::::::::
Db   2323 AATAGTAACATGTTTTCTGTATTGTTTCTAGTCCTTGGAATTATTTCTTTTATTACATTT 2382

Qy    306 PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHis 325
          |||:::||||||      :::|||:::||||||||||||||||||   |||||||||:::
Db   2383 TTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTCACTAAGCGGCTTCGATAC 2442

Qy    326 LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSer 345
          :::   |||::::::||||||   ||||||::::::||||||||||:::   :::|||:::
Db   2443 ATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTTGATGACCCTAAAAACACC 2502

Qy    346 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365
          ||||||   |||||||||   ||||||   |||   ||||||::::::||||||   |||
Db   2503 ACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAAGTTAAAGGGGCTATAGGT 2562

Qy    366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385
          ||||||:::   |||:::|||||||||      |||:::|||      :::|||||||||
Db   2563 TCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGGACAGGCATTATTATATCC 2622

Qy    386 PheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405
             ||||||||||||::::::|||   |||:::|||:::|||   |||::::::|||:::
Db   2623 TTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATTGTACCCATCATTGCAATA 2682

Qy    406 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425
             |||::::::|||         ::::::|||   |||   |||||||||:::||||||
Db   2683 GCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAGCTA 2742

Qy    426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445
          :::   ||||||||||||||||||||||||:::||||||   ||||||:::|||||||||
Db   2743 GAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTCCGAACTGTTGTTTCTTTG 2802

Qy    446 ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn 465
          |||||||||:::   ||||||   ||||||      :::   ||||||      :::||||||
Db   2803 ACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTGCAAGTACCATACAGAAAC 2862

Qy    466 ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr 485
          :::   :::||||||   |||   |||      ::::::   :::   |||      :::|||
Db   2863 TCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATCACCCAGGCAATGATGTAT 2922

Qy    486 PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr 505
          |||:::||||||         |||||||||||||||||||||:::            |||
Db   2923 TTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTGGCAAATGAGTTCATGAAC 2982

Qy    506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys 525
             :::   :::   :::||||||:::||||||   :::||||||||||||:::|||:::
Db   2983 TTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAG 3042

Qy    526 ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAla 545
                   ||||||:::|||:::|||||||||         |||||||||:::
Db   3043 GTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCAGCAGCCCACGTCATCATG 3102

Qy    546 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 565
          ::::::||||||   |||   |||||||||   |||      |||   ||||||:::|||
Db   3103 ATCATTGAAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCACGGCCTCAAGCCAAATACG 3162

Qy    566 CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspVal 585
             |||||||||:::   |||   ||||||   |||   ||||||   ||||||||||:::
Db   3163 TTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTATCCCACTCGACCAGACATC 3222

Qy    586 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605
             :::|||:::||||||||||||   :::::::::|||:::|||:::|||   ||||||
Db   3223 CCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAGACGCTGGCCCTCGTAGGT 3282

Qy    606 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProVal 625
          ||||||||||||||||||||||||   ||||||||||||:::|||   |||||||||:::
```

```
Db    3283 AGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAGCGCTTCTATGACCCCTTG 3342

Qy     626 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArg 645

Db    3343 GCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTGAATGTCCAGTGGCTCCGA 3402

Qy     646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665

Db    3403 GCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGACTGCAGCATTGCCGAGAAC 3462

Qy     666 IleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn 685

Db    3463 ATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAGATTATGCAGGCAGCCAAG 3522

Qy     686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 705

Db    3523 GAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAATACAACACCAGAGTAGGA 3582

Qy     706 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725

Db    3583 GACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATTGCCATAGCTCGCGCTCTT 3642

Qy     726 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745

Db    3643 GTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAGT 3702

Qy     746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765

Db    3703 GAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATC 3762

Qy     766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785

Db    3763 GCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTGGTGTTTCAGAATGGCAAA 3822

Qy     786 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 805

Db    3823 GTCAAGGAGCATGGCACACATCAACAGCTGCTGGCTCAGAAAGGCATCTATTTTTCCATG 3882

Qy     806 ValAsnAlaGlnSer 810

Db    3883 ATCAGTGTCCAGGCT 3897
```

RESULT 14
ABL91687
ID   ABL91687 standard; DNA; 3840 BP.
XX
AC   ABL91687;
XX
DT   28-MAY-2002 (first entry)
XX
DE   Human polynucleotide SEQ ID NO 30.
XX
KW   Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;
KW   Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;
KW   cytostatic; virucide; protozoacide; antibacterial; ds.
XX
OS   Homo sapiens.
XX
PN   DE10100586-C1.
XX
PD   11-APR-2002.
XX
PF   09-JAN-2001; 2001DE-1000586.
XX
PR   09-JAN-2001; 2001DE-1000586.
XX
PA   (RIBO-) RIBOPHARMA AG.
XX
PI   Kreutzer R, Limmer S, Rost S, Hadwiger P;
XX
DR   WPI; 2002-270454/32.
XX
PT   Inhibiting gene expression in cells, useful for e.g. treating tumors,
PT   by introducing double-stranded complementary oligoRNA having unpaired
PT   terminal bases -
XX
PS   Claim 13; Page 28-30; 104pp; German.
XX
CC   The invention relates to a method for inhibiting expression of a target
CC   gene (ABL91658-ABL91797) in a cell by introducing at least one
CC   oligoribonucleotide that has a double-stranded structure consisting of at
CC   most 49 sequential nucleotide pairs, with at least part of one strand
CC   complementary with the target gene and has at least one end a
CC   single-stranded segment of 1-4 nt. The method provides
CC   oligoribonucleotides for antisense inhibition of gene expression useful
CC   e.g. for treating tumours but the oligoribonucleotides may also be
CC   directed against genes present in pathogens (e.g. Plasmodium or
CC   viruses/viroids, pathogenic on humans, animals or plants) or against
CC   cytokine, Id, developmental or prion genes. The method provides more
CC   effective inhibition of gene expression than use of known
CC   oligonucleotides, probably because the unpaired overhang increases
CC   stability and thus intracellular concentration.
XX
SQ   Sequence 3840 BP; 1130 A; 738 C; 957 G; 1015 T; 0 other;

Alignment Scores:

Pred. No.:	5.51e-226	Length:	3840
Score:	2436.00	Matches:	466
Percent Similarity:	76.79%	Conservative:	166
Best Local Similarity:	56.62%	Mismatches:	177
Query Match:	59.72%	Indels:	14
DB:	24	Gaps:	4

US-09-873-409-2 (1-812) x ABL91687 (1-3840)

```
Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21

Db    1354 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 1413

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41

Db    1414 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 1473

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61

Db    1474 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1533

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81

Db    1534 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 1593

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101

Db    1594 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 1653

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121

Db    1654 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 1713

Qy     122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141

Db    1714 CTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGTTTGTCTACAGTT 1773

Qy     142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161

Db    1774 CGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 1833

Qy     162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178

Db    1834 GATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATGCAGACAGCAGGA 1893

Qy     179 ---------AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195

Db    1894 AATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATTGATGCCTTGGAA 1953

Qy     196 ArgLysThrAsnSerLeuProLeuHisSerVal-----------------LysSerIle 209
```

```
                   :::|||                   :::                  :::|||:::
Db    1954 ATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGTC 2013

Qy     210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228
           :::   |||          |||:::          :::   ::::::   :::   |||:::
Db    2014 CGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTGGATGAAAGTATA 2073

Qy     229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
           |||   |||||          :::|||:::||||||||         |||||||||:::   |||
Db    2074 CCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCCTTATTTTGTT 2133

Qy     249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
           :::|||          :::::::::|||||   :::   |||   |||:::||||||||
Db    2134 GTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTTGCAATAATATTT 2193

Qy     269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
           :::||||||||   :::|||          :::|||      |||   :::   ::::::
Db    2194 TCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAACGACAGAATAGT 2253

Qy     288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
              ::::::|||::::::|||:::   ||||||:::|||   |||:::::::::|||:::
Db    2254 AACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATTACATTTTTCCTT 2313

Qy     308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
           |||||         :::|||:::||||||||||||||||||   ||||||||::::::
Db    2314 CAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTT 2373

Qy     328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
           |||:::::::|||||     ||||||::::::||||||||:::      :::|||:::|||||
Db    2374 TTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAAAACACCACTGGA 2433

Qy     348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
                 ||||||||   ||||||   |||   |||||::::::||||||   ||||||||
Db    2434 GCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGG 2493

Qy     368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
           :::   |||:::|||||||||      |||:::|||      :::||||||||||||||
Db    2494 CTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATC 2553

Qy     388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
           ||||||||::::::|||   |||:::|||:::|||   |||:::::::|||:::   |||
Db    2554 TATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGA 2613

Qy     408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
           ::::::|||      ::::::|||   |||   ||||||:::||||||:::
Db    2614 GTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGT 2673

Qy     428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
           ||||||||||||||||||||||:::||||||   ||||||:::||||||||||||:::
Db    2674 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAG 2733

Qy     448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
           |||:::   ||||||   ||||||   :::   ||||||      :::||||||:::
Db    2734 GAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 2793

Qy     468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
           :::||||||   |||   |||      ::::::|||:::   |||   :::||||||:::
Db    2794 AGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCC 2853

Qy     488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
           ||||||      ||||||||||||||||||||||:::            |||:::   |||
Db    2854 TATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAG 2913

Qy     508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
              :::   :::||||||:::|||:::   :::||||||||||||:::|||:::
Db    2914 GATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGT 2973

Qy     528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
                 ||||||:::|||:::||||||||      |||||||||:::      ::::::
Db    2974 TCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATT 3033

Qy     548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
           ||||||   |||   ||||||||   |||   ||||||      |||:::|||   |||
Db    3034 GAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAA 3093

Qy     568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
           ||||||:::   |||   ||||||   |||   ||||||   |||||||||:::   :::
Db    3094 GGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTG 3153

Qy     588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
           |||:::||||||||||||   :::::::::|||:::|||:::|||   ||||||||||||
Db    3154 CTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGT 3213

Qy     608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
           ||||||||||||||||||   ||||||||||||:::|||   |||||||||:::   |||
Db    3214 GGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGG 3273

Qy     628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
           :::||||||   ||||||   :::   |||   ||||||||||||||||||||||:::
Db    3274 AAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACAC 3333

Qy     648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
           :::   ||||||   |||||||||:::||||||:::||||||||||||||||||||||
Db    3334 CTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCC 3393

Qy     668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
           ||||||||||||||||||||||||         :::||||||      ||||||   |||
Db    3394 TATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATTGTGAGGGCAGCAAAGGAGGCC 3453

Qy     688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
           |||||||||:::|||||||||      ||||||   ||||||:::|||:::||||||   |||
Db    3454 AACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAA 3513

Qy     708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
           |||   ||||||||||||||||||||||:::||||||||||||||||||||::::::
Db    3514 GGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGA 3573

Qy     728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
           :::|||   ||||||||||||||||||||||||||||||||||      :::||||||||
Db    3574 CAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAG 3633

Qy     748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
           |||||||||   ||||||||||||||||||   ||||||||||||:::|||:::   |||
Db    3634 GTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCAC 3693

Qy     768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
           |||||||||   ||||||||||||||||||||||||||||       ||||||::::::|||
Db    3694 CGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGCAGAGTCAAG 3753

Qy     788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
           |||   ||||||||||:::||||||         :::   |||||||||   :::|||:::
Db    3754 GAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGT 3813

Qy     808 AlaGlnSer 810
              |||:::
Db    3814 GTCCAGGCT 3822
```

RESULT 15
AAZ49332
ID AAZ49332 standard; cDNA; 3860 BP.
XX
AC AAZ49332;
XX
DT 14-MAR-2000 (first entry)
XX
DE Human wild-type multidrug resistance-1 (MDR-1) cDNA.
XX
KW Multidrug resistance; MDR-1; P-glycoprotein;
KW transmembrane efflux pump; haematopoietic stem cell; transduction;
KW bone marrow transplantation; chemotherapy; radiation therapy; cancer;
KW gene therapy; gene replacement; genetic defect; thalassaemia;
KW Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
KW cytokine; wild-type; ds.
XX
OS Homo sapiens.
XX

```
FH   Key             Location/Qualifiers
FT   CDS             1..3843
FT                   /*tag= a
FT                   /product= "Human wild-type MDR-1 protein"
FT   mutation        replace (553..555, GTT)
FT                   /*tag= b
FT                   /note= "cDNA sequence of G185V human mutant MDR-1 given
FT                           in AAZ49333"
XX
PN   WO9961589-A2.
XX
PD   02-DEC-1999.
XX
PF   27-MAY-1999;   99WO-US11825.
XX
PR   28-MAY-1998;   98US-0086988.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI   Sorrentino B,  Bunting K;
XX
DR   WPI; 2000-072615/06.
DR   P-PSDB; AAY58186.
XX
PT   Ex vivo expansion of hematopoietic stem cells transduced with a
PT   sequence encoding human multidrug resistance-1, used for bone marrow
PT   transplantation -
XX
PS   Claim 10; Page 68-70; 113pp; English.
XX
CC   This sequence represents cDNA encoding human wild-type
CC   multidrug resistance protein MDR-1. MDR-1 is a transmembrane
CC   efflux pump, responsible for the export of drugs from cells,
CC   particularly cancer cells. Wild-type MDR-1 shows increased
CC   resistance to etoposide and decreased resistance to vinca
CC   alkaloids compared with a mutant form (AAY58187) where the Gly at
CC   position 185 is replaced by Val. The invention relates to transducing
CC   haematopoietic stem cells with nucleic acid encoding an MDR protein
CC   and culturing the modified cells. The modified haematopoietic stem
CC   cells are useful in bone marrow transplantation (to reconstitute
CC   haematopoietic systems in patients who have undergone chemotherapy or
CC   radiation therapy) and in ex vivo gene therapy of genetic defects in
CC   cells derived from haematopoietic stem cells, e.g., thalassaemia,
CC   Gauchier's disease, sickle cell anaemia or leukaemia. The modified
CC   cells can also be used to identify factors involved in regulating
CC   proliferation and differentiation in haematopoietic stem cells.
CC   Haematopoietic stem cells that express MDR-1 will be protected against
CC   chemotherapeutic agents, so can be engrafted while the patient is
CC   undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells
CC   provides sufficient cells to permit standard biochemical analysis.
CC   Overexpression of MDR-1 allows cytokine-driven expansion of
CC   haematopoietic stem cells by at least 10-fold compared with a maximum
CC   of 4-fold in known procedures.
XX
SQ   Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;

Alignment Scores:
Pred. No.:              5.55e-226     Length:        3860
Score:                  2436.00       Matches:       466
Percent Similarity:     76.79%        Conservative:  166
Best Local Similarity:  56.62%        Mismatches:    177
Query Match:            59.72%        Indels:        14
DB:                     21            Gaps:          4

US-09-873-409-2 (1-812) x AAZ49332 (1-3860)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
              |||||    ||||||||  :::||||||||      |||:::  ||||||||
Db    1357 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 1416

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           ||||||||||||||||||||||||   |||||||||:::   ||||||:::||||||||
Db    1417 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 1476

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
           ::::::||||||   :::|||:::|||:::|||   :::||||||||||||||||||
Db    1477 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1536

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           ||||||:::   |||:::||||||:::||||||||||||||||:::|||||||||:::|||
Db    1537 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 1596

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1597 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 1656

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
           :::||||||||||||||||||||||||||||||:::||||||::::::    ||||||  |||
Db    1657 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 1716

Qy     122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
           |||:::||||||   ||||||||||||||||||||||:::||||||||||||||||||:::
Db    1717 CTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGTTTGTCTACAGTT 1776

Qy     142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
           |||:::||||||:::|||          ||||||::::::   |||||||||   |||
Db    1777 CGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 1836

Qy     162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178
              |||||||||   ::::::|||:::|||:::   ||||||      |||
Db    1837 GATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATGCAGACAGCAGGA 1896

Qy     179 ---------AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
                    :::::::::   |||   ::::::   :::|||        :::   |||
Db    1897 AATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATTGATGCCTTGGAA 1956

Qy     196 ArgLysThrAsnSerLeuProLeuHisSerVal------------------LysSerIle 209
              :::|||                                         :::|||:::
Db    1957 ATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGTC 2016

Qy     210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228
           :::   |||         |||:::            :::   ::::::   :::   |||:::
Db    2017 CGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTGGATGAAAGTATA 2076

Qy     229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
           |||   ||||||      :::|||:::|||||||||      |||||||||:::   |||
Db    2077 CCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCCTTATTTTGTT 2136

Qy     249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
           :::|||      :::::::::||||||   :::      |||   |||:::|||||||||
Db    2137 GTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTTGCAATAATATTT 2196

Qy     269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
           :::||||||   :::|||            :::|||   |||      :::   ::::::
Db    2197 TCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAACGACAGAATAGT 2256

Qy     288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
              ::::::|||::::::|||:::   ||||||:::|||   |||:::::::::|||:::
Db    2257 AACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATTACATTTTTCCTT 2316

Qy     308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
           ||||||      :::|||:::|||||||||||||||||   |||||||||::::::
Db    2317 CAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTT 2376

Qy     328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
           |||::::::||||||   ||||||::::::||||||||||:::   :::|||:::||||||
Db    2377 TTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAAAACACCACTGGA 2436

Qy     348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
              |||||||||   ||||||   |||   ||||||::::::||||||   |||||||||
Db    2437 GCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGG 2496

Qy     368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
           :::   |||:::|||||||||      |||:::|||      :::|||||||||||||||
Db    2497 CTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATC 2556
```

```
Qy      388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
            ||||||||::::::|||    |||:::|||:::|||    |||::::::|||:::    |||
Db     2557 TATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGA 2616

Qy      408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
            ::::::|||          ::::::|||   |||   |||||||||:::||||||:::
Db     2617 GTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGT 2676

Qy      428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
            ||||||||||||||||||||||||:::||||||   ||||||:::||||||||||||:::
Db     2677 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAG 2736

Qy      448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
            |||:::   ||||||   ||||||   :::   ||||||      :::||||||:::
Db     2737 GAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 2796

Qy      468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
            :::||||||   |||   |||      ::::::|||:::   |||   :::||||||:::
Db     2797 AGGAAAGCACACATCTTTGGAATTACATTTCCTTCACCCAGGCAATGATGTATTTTTCC 2856

Qy      488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
            ||||||   |||||||||||||||||||||||:::               |||:::   |||
Db     2857 TATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAG 2916

Qy      508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
               :::   :::||||||:::|||:::   :::||||||||||||:::|||:::
Db     2917 GATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGT 2976

Qy      528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
                  ||||||:::|||:::||||||||      |||||||||:::      ::::::
Db     2977 TCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATT 3036

Qy      548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
            ||||||   |||   |||||||||   |||   ||||||      |||:::|||   |||
Db     3037 GAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAA 3096

Qy      568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
            ||||||:::   |||   ||||||   |||   ||||||   |||||||||:::   :::
Db     3097 GGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTG 3156

Qy      588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
            |||:::||||||||||||   :::::::::|||:::|||:::|||   ||||||||||||
Db     3157 CTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGT 3216

Qy      608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
            ||||||||||||||||||   ||||||||||||:::|||   |||||||||:::   |||
Db     3217 GGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGG 3276

Qy      628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
            :::||||||   ||||||   :::   |||   |||||||||||||||||||||:::
Db     3277 AAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACAC 3336

Qy      648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
            :::   ||||||   |||||||||:::||||||:::||||||||||||||||||||||||
Db     3337 CTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCC 3396

Qy      668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
            ||||||||||||||||||||||||      :::||||||      ||||||      |||
Db     3397 TATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCC 3456

Qy      688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
            |||||||||:::|||||||||   ||||||   ||||||:::|||:::||||||   |||
Db     3457 AACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAA 3516

Qy      708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
            |||   |||||||||||||||||||||||||:::||||||||||||||||||||::::::
Db     3517 GGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGA 3576

Qy      728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
            :::|||   |||||||||||||||||||||||||||||||||      :::|||||||||
Db     3577 CAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAG 3636

Qy      748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
            |||||||||   |||||||||||||||||   |||||||||||||:::|||:::   |||
Db     3637 GTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCAC 3696

Qy      768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
            |||||||||   ||||||||||||||||||||||||||||         ||||||::::::|||
Db     3697 CGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGCAGAGTCAAG 3756

Qy      788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
            |||   ||||||||||||:::||||||         :::   |||||||||   :::|||:::
Db     3757 GAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGT 3816

Qy      808 AlaGlnSer 810
            |||:::
Db     3817 GTCCAGGCT 3825
```

Search completed: March 30, 2003, 03:08:04
Job time : 937.594 secs

```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 30, 2003, 03:22:44 ; Search time 5585.08 Seconds
                                           (without alignments)
                                           2354.618 Million cell updates/sec

Title:          US-09-873-409-2
Perfect score:  4079
Sequence:       1 MVDENDIRALNVRHYRDHIG..........QELLRNRDIYFKLVNAQSVQ 812

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters:      32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115420_19240/app_query.fasta_1.7544
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_30544_@runat_27032003_115420_19240 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      EST:*
                1:  em_estba:*
                2:  em_esthum:*
                3:  em_estin:*
                4:  em_estmu:*
                5:  em_estov:*
                6:  em_estpl:*
                7:  em_estro:*
                8:  em_htc:*
                9:  gb_est1:*
                10:  gb_est2:*
                11:  gb_htc:*
                12:  gb_est3:*
                13:  gb_est4:*
                14:  gb_est5:*
                15:  em_estfun:*
                16:  em_estom:*
                17:  gb_gss:*
                18:  em_gss_hum:*
                19:  em_gss_inv:*
                20:  em_gss_pln:*
                21:  em_gss_vrt:*
                22:  em_gss_fun:*
                23:  em_gss_mam:*
                24:  em_gss_mus:*
                25:  em_gss_other:*
                26:  em_gss_pro:*
                27:  em_gss_rod:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1008	24.7	943	9	AL520322	AL520322 AL520322
	2	834	20.4	998	14	BM904842	BM904842 AGENCOURT
	3	780	19.1	2676	11	AK014319	AK014319 Mus muscu
	4	770.5	18.9	1019	12	BG248052	BG248052 602359987
	5	753	18.5	726	12	BG293345	BG293345 602390738
	6	751	18.4	944	12	BF796582	BF796582 602258463
	7	743.5	18.2	886	17	BH139685	BH139685 ENTNA47TR
	8	740.5	18.2	871	17	AZ682350	AZ682350 ENTKB16TF
c	9	734.5	18.0	932	17	AZ670821	AZ670821 ENTJN69TF
	10	728.5	17.9	947	17	AZ683753	AZ683753 ENTIL96TF
	11	726.5	17.8	886	17	AZ540627	AZ540627 ENTEQ18TF
c	12	723.5	17.7	880	17	AZ687805	AZ687805 ENTLU52TF
	13	721.5	17.7	939	14	BQ720763	BQ720763 AGENCOURT
	14	720.5	17.7	897	17	AZ541090	AZ541090 ENTDS67TR
	15	711.5	17.4	913	17	BH155700	BH155700 ENTRO54TR
	16	709.5	17.4	834	17	AZ548312	AZ548312 ENTFG07TR
c	17	708.5	17.4	853	17	AZ679807	AZ679807 ENTHI67TR
	18	705.5	17.3	823	17	AZ532602	AZ532602 ENTCR50TR
	19	697.5	17.1	861	14	BQ717101	BQ717101 AGENCOURT
	20	691	16.9	780	12	BG587938	BG587938 EST489713
	21	690.5	16.9	899	17	BH154857	BH154857 ENTRO47TF
c	22	689.5	16.9	891	17	AZ682250	AZ682250 ENTKT68TR
	23	688	16.9	1123	11	AY108485	AY108485 Zea mays
	24	683.5	16.8	1341	11	AY108285	AY108285 Zea mays
	25	683	16.7	600	13	BM486593	BM486593 pgm2n.pk0
	26	682.5	16.7	919	17	AZ690701	AZ690701 ENTIY68TR
	27	678	16.6	926	14	BQ123477	BQ123477 EST609053
	28	677.5	16.6	795	12	BF313560	BF313560 601900192
	29	669	16.4	547	13	BJ030711	BJ030711 BJ030711
c	30	667	16.4	937	17	BH150760	BH150760 ENTQG42TR
c	31	663	16.3	835	13	BJ349604	BJ349604 BJ349604
	32	662	16.2	636	10	AV962688	AV962688 AV962688
c	33	661	16.2	747	13	BJ445774	BJ445774 BJ445774
	34	659.5	16.2	628	12	BG080311	BG080311 H3052B06-
	35	659	16.2	852	17	BH720383	BH720383 BOHZZ94TF
	36	657	16.1	715	14	BQ869512	BQ869512 QGD6G09.y
	37	657	16.1	785	10	AV709991	AV709991 AV709991
	38	655.5	16.1	787	12	BG584063	BG584063 EST485823
	39	655	16.1	919	14	BQ724901	BQ724901 AGENCOURT
	40	653	16.0	609	10	AV986144	AV986144 AV986144
c	41	651.5	16.0	921	17	AZ687628	AZ687628 ENTIR96TR
c	42	651	16.0	746	13	BJ380617	BJ380617 BJ380617
c	43	650	15.9	729	13	BJ353293	BJ353293 BJ353293
	44	647	15.9	1813	11	U66688	U66688 Homo sapien
c	45	641	15.7	820	13	BJ356540	BJ356540 BJ356540

ALIGNMENTS

```
RESULT 1
AL520322
LOCUS       AL520322                 943 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL520322 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB006YC15 5
            prime, mRNA sequence.
ACCESSION   AL520322
VERSION     AL520322.1  GI:12783815
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 943)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
```

```
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1..943
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DB006YC15"
                     /clone_lib="LTI_NFL004_NBC2"
                     /sex="male"
                     /tissue_type="neuroblastoma cells"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector.  Library was normalized. Library was constructed
                     by Life Technologies.  Contact : Feng Liang Life
                     Technologies, a division of Invitrogen 9800 Medical Center
                     Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
                     8371 Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      253 a    225 c    253 g    211 t     1 others
ORIGIN

Alignment Scores:
Pred. No.:              8.4e-107       Length:         943
Score:                  1008.00        Matches:        194
Percent Similarity:     80.40%         Conservative:   48
Best Local Similarity:  64.45%         Mismatches:     58
Query Match:            24.71%         Indels:         1
DB:                     9              Gaps:           0

US-09-873-409-2 (1-812) x AL520322 (1-943)

Qy    511 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 530
          ::::::|||:::|||:::   :::||||||||||||:::|||:::            |||
Db      5 TTATTATTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCT 64

Qy    531 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 550
          |||:::|||:::|||||||||      |||||||||:::      ::::::||||||
Db     65 CCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACC 124

Qy    551 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 570
          |||   |||||||||   |||   ||||||      |||:::|||   ||||||||||:::
Db    125 CCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTC 184

Qy    571 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 590
             |||   ||||||   |||   ||||||            :::|||:::|||
Db    185 ACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGA 244

Qy    591 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 610
          |||||||||   :::::::::|||:::|||:::|||   ||||||||||||||||||||||
Db    245 CTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGG 304

Qy    611 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 630
          |||||||||   ||||||||||||:::|||   ||||||||||:::   |||:::||||||
Db    305 AAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTG 364

Qy    631 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 650
             ||||||   :::   |||   |||||||||||||||||||||:::   :::   |||
Db    365 CTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATC 424

Qy    651 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 670
          |||   ||||||:::||||||:::|||||||||||||||||||||||||||||||||||||
Db    425 GTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGAC 484

Qy    671 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 690
          |||||||||||||||      :::||||||      ||||||      ||||||||||||
Db    485 AACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACAT 544

Qy    691 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 710
          :::|||||||||   ||||||   ||||||:::|||:::||||||   ||||||   |||
Db    545 GCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAG 604

Qy    711 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 730
          ||||||||||||||||||||||||:::|||||||||||||||||||||:::::::::|||
Db    605 CTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCAT 664

Qy    731 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 750
          ||||||||||||||||||||||||||||||||||||   :::||||||||||||||||||
Db    665 ATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAA 724

Qy    751 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 770
             ||||||||||||||||||   |||||||||||:::|||:::   ||||||||||||
Db    725 GAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCC 784

Qy    771 AlaIleGlnAsnAlaAspLeuIleVal-ValLeuHisAsnGlyLysIleLysGluGlnGl 790
             |||||||||||||||||||||||| |||      ||||||::::::||||||   ||
Db    785 ACCATCCAGAATGCAGACTTAATAGTGTGTGTTTCAGAATGGCAGAGTCAAGGAGCATGG 844

Qy    790 yThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSe 810
          ||||||||||:::||||||      :::   |||||||||   :::|||:::   |||::
Db    845 CACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGC 904

Qy    810 r 810
          :
Db    905 T 905

RESULT 2
BM904842
LOCUS       BM904842                 998 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6699581 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5557655
            5', mRNA sequence.
ACCESSION   BM904842
VERSION     BM904842.1  GI:19355221
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 998)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12279  row: f  column: 24
            High quality sequence stop: 738.
FEATURES             Location/Qualifiers
     source          1..998
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5557655"
                     /clone_lib="NIH_MGC_72"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      271 a    223 c    248 g    256 t
ORIGIN

Alignment Scores:
Pred. No.:              2.2e-86        Length:         998
Score:                  834.00         Matches:        164
```

```
Percent Similarity:     74.12%        Conservative:  68
Best Local Similarity:  52.40%        Mismatches:    77
Query Match:            20.45%        Indels:        4
DB:                     14            Gaps:          0

US-09-873-409-2 (1-812) x BM904842 (1-998)

Qy    335 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 354
          ||:::::||||||||:::     :::|||:::|||||     ||||||||    ||||||
Db     61 GATGTGAGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCC 120

Qy    355 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 374
             |||   ||||||::::::||||||    ||||||||:::    |||:::||||||||
Db    121 AATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAAT 180

Qy    375 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 394
                |||:::|||       :::||||||||||||||||||||||||::::::|||
Db    181 ATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTG 240

Qy    395 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 414
          |||:::|||:::|||   |||::::::|||:::   |||::::::|||          :::
Db    241 TTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTG 300

Qy    415 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 434
          :::|||   |||   |||||||||:::||||||:::   ||||||||||||||||||||||
Db    301 TCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAA 360

Qy    435 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 454
          |||:::||||||   ||||||:::||||||||||||:::|||:::   ||||||   |||
Db    361 GCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATG 420

Qy    455 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 474
          |||   :::   ||||||      :::||||||:::   :::||||||   |||   |||
Db    421 TATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGA 480

Qy    475 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 494
                ::::::|||:::   |||   :::||||||:::||||||      |||||||||
Db    481 ATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTT 540

Qy    495 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 514
          ||||||||||||:::                  |||:::   |||      :::||||||:::
Db    541 GGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCA 600

Qy    515 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 534
          |||:::   :::||||||||||||:::|||:::                ||||||:::|||:::
Db    601 GCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCC 660

Qy    535 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 554
          |||||||||      |||||||||:::      ::::::||||||   |||   ||||||
Db    661 AAAGCCAAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGAC 720

Qy    555 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 574
          |||   |||   ||||||      |||:::|||   |||||||||:::   |||
Db    721 AGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGGTGA 780

Qy    575 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer-LeuSe 594
          |||   |||   ||||||   |||||||||:::   :::|||:::|||||||| |||
Db    781 GTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCCTGGA 840

Qy    594 rIleGluArgGlyLysThrValAlaPheVal-GlySerSerGlyCysGlyLysSerThrS 614
           ::::::::|||:::|||:::|||   ||| ||||||||||||||||||||||||||
Db    841 GGTGAAGAAGGGGCAGACCCTGGCTCTGGTGGGGCAGCAGTGGCTGTGGGAAGAGCACAG 900

Qy    614 er-ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 633
             |||||||||||:::|||:::|||:::|||:::      ::::::|||   ||||||
Db    901 TGGGTCCAGCTCCTGGAGCGGGTCTACAACCCCTTGGCAAGGAAAATGCTGCTTGATGGC 960

Qy    634 ValAspAlaLys-GluLeuAsnValGlnTrp 643
                   |||    |||:::|||   |||
Db    961 AAAAGAATAAAACCGACTGGATGTTCCATGG 991

RESULT 3
AK014319
```

```
LOCUS       AK014319               2676 bp    mRNA    linear   HTC 19-JAN-2002
DEFINITION  Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length
            enriched library, clone:3222401P09:ATP-binding cassette, sub-family
            B (MDR/TAP), member 8, full insert sequence.
ACCESSION   AK014319
VERSION     AK014319.1  GI:12852089
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (strain:C57BL/6J) 14, 17 days embryo head cDNA to
            mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:3222401P09.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409 (6821), 685-690 (2001)
  MEDLINE   21085660
   PUBMED   11217851
REFERENCE   5  (bases 1 to 2676)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C.,
            Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,
            Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F.,
            Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,
            Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,
            Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
            Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,
```

```
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,
            Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and
            Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Gemomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. Second
            strand cDNA was prepared with the primer adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cleaved
            with BamHI and XhoI. cDNA of size comprised between 3 and 7 kb was
            selected before cloning. Vector: a modified pBluescript KS(+) after
            bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'
            end: BamHI. Host: DH10B.
FEATURES             Location/Qualifiers
     source          1. .2676
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:3222401P09"
                     /db_xref="MGD:MGI:1907488"
                     /db_xref="taxon:10090"
                     /clone="3222401P09"
                     /tissue_type="head"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="14, 17 days embryo"
     gene            1. .2676
                     /gene="Abcb8"
     CDS             72. .2225
                     /gene="Abcb8"
                     /note="ATP-binding cassette, sub-family B (MDR/TAP),
                     member 8
                     data source:MGD, source key:MGI:1351667, evidence:ISS
                     putative"
                     /codon_start=1
                     /protein_id="BAB29270.1"
                     /db_xref="GI:12852090"
                     /db_xref="MGD:MGI:1351667"
                     /translation="MLVHLFRFGIRGGPVPGWSLQSLRFQTFSAARSSDDRLSSHLLR
                     TVAQLRVQLRAHLPRAPPASHWSPSAWCWVGGTLVVPAVLWQHPRLCLIALCEAKESP
                     PAQPTRAPELRFNWKLFWHFLHPHLLALGAAIVLALGAALVNVQIPLLLGQLVEIVAK
                     YTRDHMGSFVSESRKLSVQLLLLYGVQGLLTFGYLVLLSHIGERMAMDMRKALFSSLL
                     RQDIAFFDAKKTGQLVSRLTTDVQEFKSSFKLVISQGLRSCTQVIGSLVSLSMLSPRL
                     TLMLAVVTPALMGVGTLMGSGLRKLSRQCQEQIARATGVADEALGNVRTVRAFAMEKR
                     EEERYQAELESCCCKAEELGRGIALFQGLSNIAFNCMVLGTLFIGGSLVAGQQLKGGD
                     LMSFLVASQTVQRSMASLSVLFGQVVRGLSAGARVFEYMALSPVIPLTGGYCIPNKDI
                     RGSITFQNVTFSYPCRPGFNVLKDFTLKLPSGKIVALVGQSGGGKTTVASLLERFYDP
                     EAGSVTLDGHDLRTLNPSWLRGQVIGFISQEPVLFATTIMENIRFGKLDASDEEVYTA
                     AREANAHEFISSFPDGYSTVVGERGTTLSGGQKQRLAIARALIKQPTVLILDEATSAL
                     DAESERVVQEALDRASAGRTVLVIAHRLSTVRAAHSIIVMANGQVCEAGTHEELLKKG
                     GLYSELIRRQTLDASLTSTPPAEKPEDPKSCQSKA"
     polyA_signal    2653. .2658
                     /gene="Abcb8"
                     /note="putative"
     polyA_site      2676
                     /gene="Abcb8"
                     /note="putative"
BASE COUNT      544 a    785 c    756 g    591 t
```

```
ORIGIN

Alignment Scores:
Pred. No.:              2.48e-79        Length:         2676
Score:                  780.00          Matches:        207
Percent Similarity:     51.28%          Conservative:   113
Best Local Similarity:  33.17%          Mismatches:     250
Query Match:            19.12%          Indels:         54
DB:                     11              Gaps:           14

US-09-873-409-2 (1-812) x AK014319 (1-2676)

Qy      216 LysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLys 235
            :::|||:::||||||      ::::::             ||||||
Db      360 GAGGCCAAAGAGTCCCCTCCTGCCCAGCCCACACGTGCCCCAGAG--------------- 404

Qy      236 IleLeuLysLeuAsn---LysProGluTrpProPhe------------------------ 246
               |||:::   |||   |||      |||   |||
Db      405 ---CTGCGCTTTAATTGGAAGCTCTTCTGGCATTTCCTTCACCCCCACCTGCTGGCCTTG 461

Qy      247 ---------ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProVal 263
                  :::||||||       |||:::::::::|||      :::   |||:::
Db      462 GGGGCAGCCATTGTGCTGGCCTTAGGTGCGGCACTAGTGAATGTGCAGATC---CCCTTG 518

Qy      264 PheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeu 283
                           :::|||:::       :::      :::            :::
Db      519 CTCCTGGGCCAGCTGGTGGAGATTGTCGCCAAGTACACGAGGGACCACATGGGGAGTTTC 578

Qy      284 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 303
                ::::::          |||:::      ::::::|||
Db      579 GTGTCTGAGTCCCGTAAGCTCAGCGTCCAGCTGCTCCTACTG------------------ 620

Qy      304 SerTyrPheMetGlnGlyLeu---------------------PheTyrGlyArgAlaGly 316
               |||   :::|||||||||                                 |||
Db      621 ---TACGGTGTTCAGGGACTGCTGACCTTTGGATACCTAGTGCTGCTGTCCCACATTGGT 677

Qy      317 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 336
            |||   :::   |||   :::|||         |||   ::::::|||   |||||||||
Db      678 GAGCGCATGGCCATGGACATGCGGAAAGCCCTTTTCAGCTCCCTGCTCCGGCAAGACATT 737

Qy      337 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 356
            |||:::||||||   |||:::      ||||||   |||   :::   |||      |||
Db      738 GCTTTCTTTGATGCCAAAAAG------ACAGGGCAGCTAGTGAGTCGCTTGACTACTGAT 791

Qy      357 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 376
            :::   :::   :::   :::               |||            ::::::   |||
Db      792 GTGCAAGAATTCAAGTCATCCTTCAAGCTTGTCATCTCCCAGGGACTGCGCAGCTGCACC 851

Qy      377 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 396
               :::      |||::::::         :::         :::|||   ::::::
Db      852 CAGGTGATTGGTAGCCTGGTGTCCCTGTCTATGCTGTCCCCTCGCCTTACCCTGATGCTG 911

Qy      397 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 416
               :::   |||   |||         |||   :::      :::   :::
Db      912 GCTGTCGTCACACCCGCCCTCATGGGAGTGGGCACCCTGATGGGCTCAGGCCTCCGAAAG 971

Qy      417 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 436
               :::   :::   ::::::::::::         |||      :::|||   |||||||||
Db      972 CTGTCTCGCCAGTGTCAGGAGCAGATTGCCAGGGCAACAGGTGTAGCAGATGAGGCCCTT 1031

Qy      437 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 456
               |||:::||||||:::   :::            ||||||      |||:::   |||:::
Db     1032 GGCAATGTTCGGACTGTGCGGGCCTTCGCGATGGAGAAGAGGGAGGAGGAACGCTATCAA 1091

Qy      457 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 476
                |||:::               :::         ||||||:::   :::|||
Db     1092 GCAGAACTGGAG------------TCATGCTGCTGTAAAGCAGAAGAACTGGGCAGGGGC 1139

Qy      477 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyr---------AlaAlaGlyPheArg 493
               |||                     |||:::                     |||
Db     1140 ATCGCCTTGTTCCAAGGGCTCTCCAACATCGCTTTCAACTGTATGGTCTTGGGCACCCTG 1199

Qy      494 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGly---------MetPhe 510
```

```
                |||            :::    |||||:::              |||          |||
Db   1200 TTCATTGGGGCTCCCTTGTGGCTGGACAGCAGCTGAAAGGGGAGACCTCATGTCCTTC 1259

Qy    511 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 530
           :::|||  :::                :::|||||                ||||||
Db   1260 CTGGTGGCTTCCCAGACAGTACAGAGGTCTATGGCC------AGCCTCTCTGTCCTGTTT 1313

Qy    531 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 550
             :::      :::        |||      |||   :::|||      :::
Db   1314 GGTCAGGTGGTACGTGGGCTGAGTGCGGGAGCCCGAGTCTTCGAATACATGGCCCTGAGC 1373

Qy    551 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 570
          |||   |||                                       |||::::::
Db   1374 CCTGTCATCCCATTGACCGGGGGCTACTGCATCCCCAACAAGGACATTCGTGGTTCCATC 1433

Qy    571 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 590
                |||:::   |||:::|||  ||||||||||||||       :::|||:::
Db   1434 ACCTTCCAAAATGTCACCTTCAGCTACCCCTGCAGACCTGGCTTCAATGTGCTCAAGGAC 1493

Qy    591 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 610
             :::|||   :::        ||||||   ||||||   ||||||   ||||||   |||
Db   1494 TTCACCCTGAAGCTGCCCTCTGGCAAGATTGTGGCTCTTGTGGGCCAGTCTGGGGGAGGA 1553

Qy    611 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 630
          |||:::|||      ||||||:::|||  |||||||||      |||   |||
Db   1554 AAGACCACAGTTGCCTCCTTGCTGGAACGCTTCTATGACCCTGAAGCTGGCTCGGTGACG 1613

Qy    631 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln---IleAla 649
                ||||||   |||   :::   ||||||      ||||||||   |||   |||
Db   1614 TTGGATGGGCATGACCTGCGAACTCTCAACCCCTCCTGGCTCCGGGGCCAGGTCATAGGT 1673

Qy    650 IleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGly 669
             :::   ||||||||||||||||||      :::|||   |||||||||   :::|||
Db   1674 TTCATCAGCCAGGAGCCAGTCCTGTTTGCAACAACCATCATGGAGAATATCCGATTTGGG 1733

Qy    670 ------AspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
                |||   |||                :::|||:::      ||||||      |||
Db   1734 AAGCTGGATGCTTCCGAT------------GAAGAGGTGTACACAGCTGCACGAGAAGCC 1781

Qy    688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
          |||   |||   ||||||            |||:::   |||:::|||   ||||||   :::
Db   1782 AATGCCCACGAGTTCATCAGCAGCTTCCCCGATGGCTACAGCACTGTGGTTGGTGAGCGG 1841

Qy    708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
          |||      ||||||||||||||||||||||||||||||||||||||||||||||||::::::
Db   1842 GGCACAACCTTGTCTGGTGGCCAGAAGCAGCGCCTAGCCATCGCACGTGCCCTCATCAAG 1901

Qy    728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
          :::|||   :::|||:::||||||||||||||||||||||||||||   :::||||||:::
Db   1902 CAGCCCACAGTGCTGATCCTGGACGAGGCCACCAGTGCGCTAGATGCAGAATCCGAGAGG 1961

Qy    748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
          ||||||||   |||||||||:::|||      |||||||||   ||||||:::   |||
Db   1962 GTGGTACAGGAGGCCCTGGACCGGGCCAGTGCTGGCCGCACCGTGTTGGTCATTGCCCAC 2021

Qy    768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
          |||||||||   ::::::   |||      |||:::|||:::   ||||||::::::
Db   2022 CGGCTTAGTACTGTCCGTGCAGCCCACTCCATCATTGTCATGGCCAATGGCCAAGTCTGT 2081

Qy    788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
          |||   |||||||||:::|||||||||:::               :::|||   :::|||:::
Db   2082 GAGGCTGGGACCCACGAAGAACTCCTTAAAAAAGGCGGGCTCTATTCAGAGCTTATCCGG 2141

Qy    808 AlaGlnSerVal 811
                |||::::::
Db   2142 AGACAAACCCTG 2153
```

RESULT 4
BG248052

```
LOCUS       BG248052                1019 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  602359987F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4488404 5',
            mRNA sequence.
ACCESSION   BG248052
VERSION     BG248052.1  GI:12757867
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1019)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10334  row: n  column: 21
            High quality sequence stop: 650.
FEATURES             Location/Qualifiers
     source          1..1019
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4488404"
                     /clone_lib="NCI_CGAP_Mam1"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="10 months, virgin"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
BASE COUNT      276 a    262 c    283 g    196 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:               6.2e-79        Length:         1019
Score:                   770.50         Matches:        154
Percent Similarity:      79.17%         Conservative:   36
Best Local Similarity:   64.17%         Mismatches:     47
Query Match:             18.89%         Indels:         3
DB:                      12             Gaps:           1

US-09-873-409-2 (1-812) x BG248052 (1-1019)

Qy    567 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 586
          ||||||||::::::|||         |||   |||   ||||||   ||||||   :::
Db      4 GAAGGAAATGTGCAATTTAGTGGAGTCGTGTTCAACTATCCCACCCGACCCAGCATCCCA 63

Qy    587 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 606
          :::|||:::||||||||||||||   :::::::::|||:::|||:::|||   |||||||||
Db     64 GTGCTTCAGGGGCTGAGCCTTGAGGTGAAGAAGGGCCAGACGCTGGCCCTGGTGGGCAGC 123

Qy    607 SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGln 626
          ||||||||||||||||||||   ||||||||||||:::|||   |||||||||:::
Db    124 AGTGGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAGCGCTTCTACGACCCCATGGCT 183

Qy    627 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 646
             |||   |||      ||||||   :::   |||:::||||||||||||||||||||:::
Db    184 GGATCAGTGTTTCTAGATGGCAAAGAAATAAAGCAACTGAATGTCCAGTGGCTCCGAGCA 243

Qy    647 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 666
          |||:::   ||||||   ||||||||:::||||||:::||||||||||||||||||||
Db    244 CAGCTGGGCATTGTGTCCCAAGAGCCCATTCTCTTTGACTGCAGCATCGCAGAGAACATT 303

Qy    667 AlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla 686
          ||||||||||||||||||||||||               :::||||||   ||||||
Db    304 GCCTACGGAGACAACAGCCGGGTCGTGTCTTATGAGGAGATTGTGAGGGCAGCCAAGGAG 363

Qy    687 AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeu 706
```

```
             |||||||||||  ||||||:::  ||||||:::||||||||||||:::||||||
Db     364 GCCAACATCCACCAGTTCATCGACTCGCTACCTGATAAATACAACACCAGAGTAGGAGAC 423

Qy     707 LysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeu 726
           |||||  |||||||||||||||||||||||||||:::||||||||||||||||||:::
Db     424 AAAGGCACTCAGCTGTCGGGTGGGCAGAAGCAGCGCATCGCCATCGCACGCGCCCTCGTC 483

Qy     727 GlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGlu 746
           ::::::|||  |||||||||||||||||||||||||||||||||||    :::||||||
Db     484 AGACAGCCTCACATTTTACTTCTGGACGAAGCAACATCAGCTCTGGATACAGAAAGTGAA 543

Qy     747 LysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThr 766
           |||||||||||  ||||||||||||||||||    ||||||||||:::|||:::
Db     544 AAGGTTGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACCTGCATTGTGATCGCT 603

Qy     767 HisArgLeuSerAlaIleGlnAsnAlaAspLeuIle-ValValLeuHis-AsnGlyLysI 786
           |||||            |||||||||||||||| ||||||:::  |||  |||:
Db     604 CACCGCTGTCCACCATCCCAGAACGCGGACTTGATCGGTGGTGATTCAAGAACCGGAAGG 663

Qy     786 leLys---GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 803
           ::|||      :::        |||:::::::::      :::   ||||||||
Db     664 TCAAGGGAGCACGGGCACCCCAACAGCAAGTGGTGGCGCAGAAGGGCATCTACTTC 719
```

```
RESULT 5
BG293345
LOCUS       BG293345                726 bp    mRNA    linear   EST 21-FEB-2001
DEFINITION  602390738F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4502552 5',
            mRNA sequence.
ACCESSION   BG293345
VERSION     BG293345.1  GI:13052943
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 726)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10371  row: l  column: 09
            High quality sequence stop: 658.
FEATURES             Location/Qualifiers
     source          1..726
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4502552"
                     /clone_lib="NIH_MGC_94"
                     /tissue_type="retina"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 3.3 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      189 a    220 c    190 g    127 t
ORIGIN

Alignment Scores:
Pred. No.:              3.9e-77      Length:         726
Score:                  753.00       Matches:        152
Percent Similarity:     84.29%       Conservative:   25
Best Local Similarity:  72.38%       Mismatches:     30
Query Match:            18.46%       Indels:         3
DB:                     12           Gaps:           0
```

```
US-09-873-409-2 (1-812) x BG293345 (1-726)

Qy     601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 620
           :::|||  ||||||||||||||||||||||||||||||    ||||||||||:::|||
Db       2 CTGGCCCTGGTGGGCAGCAGTGGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAGCGC 61

Qy     621 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 640
           |||||||||:::  |||  ||||||  ||||||    :::||||||:::||||||
Db      62 TTCTATGACCCCATGGCTGGATCAGTGCTCTTAGATGGTCAAGAAGCAAAGAAACTCAAT 121

Qy     641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
           |||||||||||||:::|||:::  ||||||  |||||||||:::||||||:::|||
Db     122 GTCCAGTGGCTCCGAGCTCAACTGGGCATTGTGTCCCAGGAACCCATTCTCTTTGACTGC 181

Qy     661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 680
           ||||||||||||||||||||||||||||||||||||||||||||||  |||||||||
Db     182 AGCATCGCAGAGAACATCGCCTATGGAGACAACAGCCGGGTCGTGCCTCATGATGAGATT 241

Qy     681 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700
                 ||||||    ||||||||||||  |||||||||  ||||||:::||||||
Db     242 GTGAGGGCAGCCAAGGAGGCCAACATCCACCCCTTCATCGAGACGCTGCCCCAAAAATAT 301

Qy     701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720
           ||||||:::||||||  ||||||  ||||||||||||||||||||||||:::|||
Db     302 AACACAAGAGTAGGAGACAAGGGGACGCAGCTCTCTGGGGGCCAGAAGCAGAGGATTGCC 361

Qy     721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740
           ||||||||||||||:::::::::|||::::::|||||||||||||||||||||||||
Db     362 ATCGCCCGAGCCCTCATCAGACAGCCTCGGGTCCTACTGCTGGATGAAGCCACGTCAGCT 421

Qy     741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
           ||||||  :::||||||||||||||||||  |||||||||||||||||  ||||||
Db     422 CTGGATACTGAGAGTGAAAAGGTTGTCCAGGAAGCACTGGACAAAGCCAGGGAAGGCCGC 481

Qy     761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla-AspLeuIleValVa 780
           ||||||:::  :::  ||||||||||||  ||||||||||||   |||||||||||
Db     482 ACCTGCATTGCGATCGCTCACCGCCTGTCCACCATCCAGAACGCGGCACTTGATCGTGGT 541

Qy     780 lLeuHisAsnGlyLysIleLysGluGlnGlyThr-HisGlnGluLeuLeuArgAsnArgA 800
           |:::  |||||||||:::||||||  ||||||  ||||||:::||||||      :::
Db     542 GATTGAGAACGGCAAGGTCAAGGAGCACGGCACCCCACCAGCAGCTGCTGGCGCAGAAGG 601

Qy     800 spIleTyrPhe-LysLeuValAsn 807
            ||||||||| :::||||||||||
Db     602 GCATCTATTTCTCAACTGGTCAAC 625
```

```
RESULT 6
BF796582
LOCUS       BF796582                944 bp    mRNA    linear   EST 12-JAN-2001
DEFINITION  602258463F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4341710 5',
            mRNA sequence.
ACCESSION   BF796582
VERSION     BF796582.1  GI:12101636
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 944)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis Staudt, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9955  row: n  column: 15
```

```
            High quality sequence start: 9
            High quality sequence stop: 669.
FEATURES             Location/Qualifiers
     source          1..944
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4341710"
                     /clone_lib="NIH_MGC_85"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.867 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      265 a    241 c    237 g    201 t
ORIGIN

Alignment Scores:
Pred. No.:              1.05e-76      Length:         944
Score:                  751.00        Matches:        166
Percent Similarity:     74.62%        Conservative:   31
Best Local Similarity:  62.88%        Mismatches:     60
Query Match:            18.41%        Indels:         9
DB:                     12            Gaps:           2

US-09-873-409-2 (1-812) x BF796582 (1-944)

Qy    533 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 552
          |||:::||||||||      ||||||||||||||    |||    |||:::::::|||
Db     39 TATGCTAAAGCTAAGCTGTCTGCAGCCCACTTATTCATGCTGTTGAAAGACAACCTCTG 98

Qy    553 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 572
          |||||||||   |||:::||||||   |||||||||         ||||||||||:::   |||
Db     99 ATTGACAGCTACAGTGAAGAGGGGCTGAAGCCTGATAAATTTGAAGGAAATATAACATTT 158

Qy    573 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 592
                ||||||   |||   ||||||    |||      :::||||   :::||||||||||
Db    159 AATGAAGTCGTGTTCAACTATCCCACCCGAGCAAACGTGCCAGTGCTTCAGGGGCTGAGC 218

Qy    593 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 612
          |||    :::::::::|||:::|||:::|||   ||||||||||||||||||||||||||||
Db    219 CTGGAGGTGAAGAAAGGCCAGACACTAGCCCTGGTGGGCAGCAGTGGCTGTGGGAAGAGC 278

Qy    613 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 632
          |||    ||||||||||||:::|||   ||||||||||:::    |||    ||||||    |||
Db    279 ACGGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCGGGGACAGTGCTTCTCGAT 338

Qy    633 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 652
          |||    :::||||||:::||||||||||||||||||||||:::|||:::      ||||||
Db    339 GGTCAAGAAGCAAAGAAACTCAATGTCCAGTGGCTCAGAGCTCAACTCGGAATCGTGTCT 398

Qy    653 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 672
          |||||||||:::||||||:::||||||||||||||||||||||||||||||||||||||||
Db    399 CAGGAGCCTATCCTATTTGACTGCAGCATTGCCGAGAATATTGCCTATGGAGACAACAGC 458

Qy    673 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 692
          |||||||||      |||||||||      ||||||    |||||||||||||||    |||
Db    459 CGGGTTGTATCACAGGATGAAATCGTGAGTGCAGCCAAAGCTGCCAACATACATCCTTTC 518

Qy    693 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 712
          ||||||   ||||||   ||||||   |||:::||||||   ||||||    |||||||||
Db    519 ATCGAGACGTTACCCCACAAATATGAAACAAGAGTGGGAGATAAGGGGACTCAGCTCTCA 578

Qy    713 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLys-ProLysIleLe 732
          ||||||    ||||||||:::||||||||||||||||||:::::::::: |||:::||||||
Db    579 GGAGGTCACAAACAGAGGATTGCTATTGCCCGAGCCCTCATCAGACAACCCTCAAATCCT 638

Qy    732 uLeuLeuAspGluAlaThrSerAla-LeuAspAsnAspSerGluLysValVal------G 750
          ||||||||||||||||||||||||| ||||||             |||              :
Db    639 CCTGTTGGATGAAGCTACATCAGCCTCTGGATACCTGACAAGTGAACACGGTTGTCCCAA 698

Qy    750 lnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuS 770
          ::                |||                ||||||||||||||:::     ||||||
Db    699 GAAGCCCCTGTCACACAGCAGAAGAGCGCCGCACCTGCCTTGTGATTGCTCACCGCTGTC 758

Qy    770 erAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnG 790
              ||||||||||||||   |||||||||              |||::::::|||
Db    759 AAC--ATCCAGAATGCAGACTT-ATAGTGGTGTTCCCAACGGGAGAAGTCAAG------- 808

Qy    790 lyThrHis 792
             ||||||
Db    809 --ACACAT 814

RESULT 7
BH139685
LOCUS       BH139685                 886 bp    DNA     linear   GSS 07-AUG-2001
DEFINITION  ENTNA47TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   BH139685
VERSION     BH139685.1  GI:15098746
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
  ORGANISM  Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 886)
  AUTHORS   Loftus,B., Wang,Z., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library (2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Reverse
            Class: shotgun
            High quality sequence start: 18
            High quality sequence stop: 851.
FEATURES             Location/Qualifiers
     source          1..886
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a  method described by Clark and Diamond (Clark,
                     C.G., and Diamond, L.S.  (1993) Entamoeba histolytica: a
                     method for isolate identification.  Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + i method used for
                     the library  construction is described in detail in Smith,
                     H.O. and Venter, J.C.  (Making small insert libraries for
                     whole genome shotgun sequencing  projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin  and B.
                     Barell, Oxford University Press, 1999)."
BASE COUNT      348 a    105 c    176 g    257 t
ORIGIN

Alignment Scores:
Pred. No.:              7.1e-76       Length:         886
Score:                  743.50        Matches:        151
Percent Similarity:     73.43%        Conservative:   48
Best Local Similarity:  55.72%        Mismatches:     71
Query Match:            18.23%        Indels:         2
DB:                     17            Gaps:           1

US-09-873-409-2 (1-812) x BH139685 (1-886)

Qy    540 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 559
```

```
               |||    :::    :::    :::::::::::|||||    ||||||       |||:::|||
Db        20 GCAAAGAATACTTATGATGTTATTGATAGAAAACCACT-ATTGATTGTTATAGTGAAGAA 78

Qy       560 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 579
             |||:::       :::      :::|||    :::||||||:::::::::   |||   |||
Db        79 GGTGAAACATTTAATGATGTTAAAGGTGAAATTGAATTTAAAGACATTTGTTTCAGATAT 138

Qy       580 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 599
             |||   |||||||||       :::|||:::|||:::|||       :::|||:::||||||
Db       139 CCAACAAGACCAGACAATTCTGTCTTGAAAGGTATTTCATTCAAAGTAGAACAAGGAAAA 198

Qy       600 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 619
             |||||||||      ||||||:::|||||||||||||||||||||||||||||||||:::::
Db       199 ACTGTTGCATTAGTAGGAGCATCAGGATGTGGTAAATCAACATCAGTTCAGTTGATTGAA 258

Qy       620 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 639
             |||   |||||||||      |||      ||||||   ||||||    :::   |||:::|||
Db       259 AGATTTTATGATCCAACACATGGAGATGTATTATTAGATGGACATAATATCAAAGATTTG 318

Qy       640 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 659
             |||:::    :::||||||||||||||||    :::|||   |||||||||||||||||
Db       319 AATATTCATTTCTTAAGAAGTCAAATTGGAATGGTAGGACAAGAACCAGTATTATTTGCT 378

Qy       660 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal---ValProLeuAsp 678
                |||:::    :::||||||         |||             |||   |||       :::
Db       379 GAAAGTGTTATGGATAATATTAGAAGAGGAGTACCTAAAGGAGTTGAAGTAAGTAATGAA 438

Qy       679 GluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGlu 698
             :::|||       ||||||    ||||||   |||   ||||||       :::||||||
Db       439 CAAATTTATGCTGCTGCTAAAATGGCAAATGCACATGACTTTATTTCAGCAATGCCAGAA 498

Qy       699 LysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArg 718
                |||||||||   ||||||    :::|||||||||:::|||||||||||||||||||||
Db       499 GGATATAACACAATGGTAGGTGATAGAGGTGCACAAATTTCAGGAGGACAAAAACAAAGA 558

Qy       719 LeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThr 738
             :::|||||||||||||||||||:::::   ||||||:::|||||||||||||||||||
Db       559 ATTGCTATTGCACGTGCATTGATTAGAAATCCAAAAGTGTTATTACTCGATGAAGCTACA 618

Qy       739 SerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThr 758
             ||||||||||||:::::|||||||||:::||||||   |||||||||||||||
Db       619 TCAGCACTTGATTCAGAAAGTGAAAAGATTGTACAAGATGCACTTGACAAAGCAGCAAAA 678

Qy       759 GlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIle 778
             |||||||||    :::|||:::   ||||||||||||||||||||   |||
Db       679 GGAAGAACAACAATTGTAATTGCACATAGATTATCAACTATTCAAAATGCAGATCAAATA 738

Qy       779 ValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsn 798
                |||:::      |||:::|||   |||:::||||||||||||||||||||||
Db       739 TGTGTTATTATGAGAGGAAGAATTGCAGAAAGAGGAACACATCAAGAGTTATTAGATTTG 798

Qy       799 ArgAspIleTyrPheLysLeuValAsnAlaGln 809
             :::       |||:::    |||          |||
Db       799 AAAGGATTTTATTATACACTTGCTATGCAACAA 831
```

```
RESULT 8
AZ682350
LOCUS       AZ682350                871 bp    DNA     linear   GSS 14-DEC-2000
DEFINITION  ENTKB16TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   AZ682350
VERSION     AZ682350.1  GI:11819496
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
  ORGANISM  Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 871)
  AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Forward
            Class: shotgun
            High quality sequence start: 16
            High quality sequence stop: 860.
FEATURES             Location/Qualifiers
     source          1..871
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a  method described by Clark and Diamond (Clark,
                     C.G., and Diamond, L.S.  (1993) Entamoeba histolytica: a
                     method for isolate identification.  Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + i method used for
                     the library  construction is described in detail in Smith,
                     H.O. and Venter, J.C.  (Making small insert libraries for
                     whole genome shotgun sequencing  projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin  and B.
                     Barell, Oxford University Press, 1999)."
BASE COUNT      349 a    117 c    174 g    231 t
ORIGIN

Alignment Scores:
Pred. No.:                1.55e-75       Length:        871
Score:                    740.50         Matches:       147
Percent Similarity:       70.04%         Conservative:  47
Best Local Similarity:    53.07%         Mismatches:    82
Query Match:              18.15%         Indels:        1
DB:                       17             Gaps:          1

US-09-873-409-2 (1-812) x AZ682350 (1-871)

Qy       534 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 553
             :::       ||||||         |||       :::::::::      :::::::::    |||:::|||
Db        39 AATAGTGCTAAAGTTGCCGCATTTAATGTTTATCAAACAATTGATAGAATCCCAGATATT 98

Qy       554 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 573
             |||    :::|||         |||:::    |||         |||    ||||||:::    |||
Db        99 GATTGTCAGTCTATTGGAGGTGAATGTCCAACTGAGTGTAATGGAAATATTAGATTTGAG 158

Qy       574 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 593
             :::|||   |||   ||||||    |||                   :::|||:::||||||    |||
Db       159 GATGTTCAATTTGTTTATCCAACAAGACTGTCTCATCATGTATTAAAAGGACTTGACCTT 218

Qy       594 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 613
                |||::::::|||:::|||:::|||      ||||||:::|||||||||||||||||||||
Db       219 GAAATTAAGAAAGGACAAACAATTGCATTAGTTGGAGCATCAGGATGTGGGAAGTCAACT 278

Qy       614 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 633
             ::::::||||||:::||||||   |||||||||      |||:::|||         ||||||
Db       279 ACTATTCAATTAATCCAAAGAAATTATGATCCAAATGGTGGAAGAGTAACATTAGACGGA 338

Qy       634 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 653
                |||    :::|||||||||::::::|||||||||:::||||||      :::|||   |||
Db       339 AAAGATATACGAGAGTTGAATATCAAATGGTTAAGAAATCAAATAGGATTAGTTGGACAA 398

Qy       654 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 673
             |||||||||||||||             :::|||   |||||||||         |||
Db       399 GAACCAGTGTTGTTTGCAGGAACAATTCGAGAAAATATTATGCTTGGAGCTAAAGAAGGA 458

Qy       674 ValValProLeu---AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 692
                   |||         :::|||:::    |||    |||          ||||||    |||    |||
```

```
Db     459 GCAACACCAAGTGAAGAAGAGATGATTGAATGTGCTAAAATGGCAAATGCACATGACTTC 518

Qy     693 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 712
           |||      ||||||||   |||:::|||   :::|||  ||||||||  ||||||
Db     519 ATTTCTAAACTTCCAGAAGGATATGACACAATAATTGGAGAAAAAGGAGCATTATTATCA 578

Qy     713 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 732
           |||||||||||||||||||:::|||||||||||||||||:::::    |||   ||||||
Db     579 GGAGGACAAAAACAAAGAATTGCAATTGCACGTGCATTGATTCGAAACCCATCTATTCTT 638

Qy     733 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 752
           |||||||||||||||||||||||||||||||       ||||||||:::||||||   |||
Db     639 CTTCTTGATGAAGCTACATCAGCACTTGATACACAAAGTGAAAAGATTGTACAAGAAGCA 698

Qy     753 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 772
           |||:::||||||       ||||||||   ::::::|||  ||||||||:::    :::
Db     699 CTTGAAAAAGCATCTAAAGGAAGAACAACAATTATTGTAGCACATAGACTAACAACTGTT 758

Qy     773 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 792
           :::||||||||   |||   |||   |||   |||:::|||   ||||||||   |||
Db     759 AGAAATGCAGATAAAATATGTGTATTTCATCAAGGAGAAATTATTGAACAAGGAAAACAT 818

Qy     793 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGln 809
           ||||||||:::        :::      |||:::   ||||||      |||
Db     819 CAAGAATTAATGGATTTGAAAGGAACATATTATGGATTAGTCAAAAGACAA 869

RESULT 9
AZ670821/c
LOCUS       AZ670821                 932 bp    DNA     linear   GSS 14-DEC-2000
DEFINITION  ENTJN69TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   AZ670821
VERSION     AZ670821.1  GI:11807967
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
  ORGANISM  Entamoeba histolytica.
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 932)
  AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Forward
            Class: shotgun
            High quality sequence start: 46
            High quality sequence stop: 867.
FEATURES             Location/Qualifiers
     source          1..932
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a  method described by Clark and Diamond (Clark,
                     C.G., and Diamond, L.S.  (1993) Entamoeba histolytica: a
                     method for isolate identification.  Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + i method used for
                     the library  construction is described in detail in Smith,
                     H.O. and Venter, J.C.  (Making small insert libraries for
                     whole genome shotgun sequencing  projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin  and B.
                     Barell, Oxford University Press, 1999)."
BASE COUNT      258 a    182 c    125 g    367 t
ORIGIN

Alignment Scores:
Pred. No.:               8.76e-75      Length:          932
Score:                   734.50        Matches:         150
Percent Similarity:      71.11%        Conservative:    42
Best Local Similarity:   55.56%        Mismatches:      77
Query Match:             18.01%        Indels:          2
DB:                      17            Gaps:            1

US-09-873-409-2 (1-812) x AZ670821 (1-932)

Qy     533 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 552
           ::::::   ||||||:::   |||      ::::::      :::::::::   |||:::
Db     835 TTTGCTACAGCTAAAGCTTCTGCATATAGAATTTATCAAACAATTGATAGAATCCCAGAT 776

Qy     553 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 572
           ||||||  ||||||         |||:::   |||         |||   ||||||:::
Db     775 ATTGATTGTAGATCTACAGCTGGTGAATGTCCAACTGAGTGTAATGGAAATATTACATTA 716

Qy     573 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 592
              :::|||   |||   ||||||   ||||||            ||||||   ||||||
Db     715 GAAGATGTTCAATTTAGATATCCAACAAGACCAACTAAACAAATTCTTGGTGGACTTGAT 656

Qy     593 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 612
           |||   |||:::::|||:::||||||||||||   |||||||||||||||||||||||||
Db     655 CTTGAAATTAAGAAAGGACAAACAGTTGCATTAGTAGGAGCATCAGGATGTGGTAAATCA 596

Qy     613 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 632
           |||:::::::||||||:::||||||   ||||||||||||   |||   |||      |||
Db     595 ACTACTATTCAATTAGTCCAAAGAAATTATGATCCAGTTGGTGGGTCAGTCAAATTAGAT 536

Qy     633 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 652
           |||   |||   ::::::||||||:::::::|||||||||:::||||||    :::|||
Db     535 GGAAAAGATTTAAGAGATTTAAATATCAAATGGTTAAGAAATCAAATAGGATTAGTTGGA 476

Qy     653 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 672
           |||||||||:::||||||   |||:::|||   |||||||||         |||
Db     475 CAAGAACCTATTTTGTTTGCATGTACTATTAGAGAAAATATTATGCTTGGAGCTAGAGAT 416

Qy     673 ArgValValProLeu---AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSer 691
                       |||      :::|||:::   |||   |||      ||||||   |||
Db     415 GGAGAAACACCAACTGAAGAAGAGATGATTGAATGTGCTAAAATGGCAAATGCACATGAA 356

Qy     692 PheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeu 711
           ||||||      |||||||||   |||:::|||   ||||||   |||||||||   |||
Db     355 TTTATTTCTCATCTTCCAGAAGGATATGATACAATGGTAGGAGAAAAAGGAGCTGCATTA 296

Qy     712 SerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIle 731
           |||||||||||||||||||||:::||||||||||||||||||||:::::::||||||   |||
Db     295 TCAGGAGGACAAAAACAAAGAATTGCTATTGCACGTGCATTGATTAGAAAACCTACAATT 236

Qy     732 LeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHis 751
           |||||||||||||||||||||||||||||||        |||||||||:::||||||
Db     235 TTATTACTTGATGAAGCTACATCAGCACTTGATACACAAAGTGAAAAGATTGTACAACAA 176

Qy     752 AlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAla 771
           ||||||:::||||||       |||||||||   :::||||||   |||||||||:::
Db     175 GCACTTGAAAAAGCTAGTCAAGGAAGAACAACAATTGTTGTAGCACATAGATTAACAACT 116

Qy     772 IleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThr 791
           :::::||||||   |||   |||   |||   |||:::|||   ||||||||||||
Db     115 GTTAGAAATGCAAGTAGAATTTGTGTATTCCATCAAGGAGAAATTAT-GAACAAGGAACA 57

Qy     792 HisGlnGluLeuLeuArgAsnArgAspIle 801
           ||||||||||||      :::|||:::
Db      56 CATCAAGAATTAAGGATTTAAAAGGATTTA 27

RESULT 10
AZ683753
LOCUS       AZ683753                 947 bp    DNA     linear   GSS 14-DEC-2000
```

```
DEFINITION  ENTIL96TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   AZ683753
VERSION     AZ683753.1  GI:11820899
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
  ORGANISM  Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 947)
  AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Forward
            Class: shotgun
            High quality sequence start: 40
            High quality sequence stop: 778.
FEATURES             Location/Qualifiers
     source          1..947
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a  method described by Clark and Diamond (Clark,
                     C.G., and Diamond, L.S.  (1993) Entamoeba histolytica: a
                     method for isolate identification.  Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + i method used for
                     the library  construction is described in detail in Smith,
                     H.O. and Venter, J.C.  (Making small insert libraries for
                     whole genome shotgun sequencing  projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin  and B.
                     Barell, Oxford University Press, 1999)."
BASE COUNT      379 a    121 c    178 g    269 t
ORIGIN

Alignment Scores:
Pred. No.:              4.54e-74      Length:          947
Score:                  728.50        Matches:         149
Percent Similarity:     69.82%        Conservative:    43
Best Local Similarity:  54.18%        Mismatches:      81
Query Match:            17.86%        Indels:          2
DB:                     17            Gaps:            1

US-09-873-409-2 (1-812) x AZ683753 (1-947)

Qy    536 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSer 555
          ||||||:::   |||      :::|||      :::::::::   |||:::||||||
Db     69 GCAAAAGCATCAGCCTATAAAATATTTACAACAATTGATAGAATCCCAGATATTGATTGT 128

Qy    556 ArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 575
          :::|||      |||:::   |||   ||||||:::   |||   |||   :::|||
Db    129 CAGTCTATTGGAGGTGAATGTCCAACTGAGTGTAATGGGAATATTAGATTTGATGATGTT 188

Qy    576 SerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIle 595
             |||  ||||||  ||||||              :::|||:::|||:::   |||   |||
Db    189 CAATTTGTTTATCCAACAAGACCATCTCATCATGTATTAAAAGGAATTGACCTTGAAATT 248

Qy    596 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 615
          ::::::|||:::   :::|||   ||||||:::|||||||||   |||||||||::::::
Db    249 AAGAAAGGAGAAATAATTGCATTAGTTGGAGCATCAGGATGTAGAAAGTCAACTACTATT 308

Qy    616 GlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAsp 635
          ||||||:::||||||   |||:::|||      |||:::|||      ||||||   |||
Db    309 CAATTAATCCAAAGAAATTATGAACCAAATGGTGGAAGAGTAACATTAGATGGGAAAGAC 368

Qy    636 AlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 655
             :::|||||||||::::::|||||||||:::||||||   ||||||   |||||||||
Db    369 ATAAGAGAGTTGAATATCAAATGGTTAAGAAATCAAATAGGAATAGTTGGACAAGAACCA 428

Qy    656 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGly---AspAsnSerArgVal 674
          |||||||||:::   :::|||   |||||||||         |||
Db    429 GTATTATTTTCAGGAACAATTGGAGAAAATATCATGCTTGGAGCTAAAGAAGGAGAAACA 488

Qy    675 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 694
          :::      :::|||:::   |||            ||||||   |||   ||||||
Db    489 CTAAGTAAAGAAGAGATGATTGAATGTACTAAAATGGCAAATGCACATGACTTTATTTCT 548

Qy    695 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 714
             |||||||||   |||:::|||   :::|||   |||||||||   ||||||||||||
Db    549 AAACTTCCAGAAGGATATGACACAATAATTGGAGAAAAAGGAGCATTATTATCAGGAGGA 608

Qy    715 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 734
          ||||||||||||:::||||||||||||||||||::::::   |||   ||||||||||||
Db    609 CAAAAACAAAGAATTGCAATTGCACGTGCATTGATTCGAAACCCATCTATTCTTCTTCTT 668

Qy    735 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 754
          ||||||||||||||||||||||||      |||||||||:::||||||   |||||||||
Db    669 GATGAAGCTACATCAGCACTTGATACACAAAGTGAAAAGATTGTACAAGAAGCACTTGAC 728

Qy    755 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 774
          ||||||      |||||||||   ::::::|||   ||||||||||||   |||||||||
Db    729 AAAGCATCAAAAGGAAGAACAACAATTATTGTAGGACATAGACTATCGACTATTCAAAAT 788

Qy    775 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 794
          ||||||   |||   |||:::      |||||||||   ||||||||||||||||||||||
Db    789 GCAGATCAAATATGTGTTATTATGAGAGGAAAAATAGTAGAACAAGGAACACATCAAGAA 848

Qy    795 -LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 808
           ||||||      :::      |||:::   |||      |||
Db    849 TTTATTGGATTTGAAAAGATTTATTATACACTTGCTTATGCA 891

RESULT 11
AZ540627
LOCUS       AZ540627              886 bp    DNA     linear   GSS 14-NOV-2000
DEFINITION  ENTEQ18TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   AZ540627
VERSION     AZ540627.1  GI:11147603
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
  ORGANISM  Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 886)
  AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Forward
            Class: shotgun
            High quality sequence start: 23
            High quality sequence stop: 856.
FEATURES             Location/Qualifiers
     source          1..886
                     /organism="Entamoeba histolytica"
```

```
                    /strain="HM1:IMSS"
                    /db_xref="taxon:5759"
                    /clone_lib="Entamoeba histolytica Sheared DNA"
                    /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                    Institute for Genomic Research (TIGR), Rockville, MD.
                    Genomic DNA isolated from broth cultures of E. histolytica
                    using a  method described by Clark and Diamond (Clark,
                    C.G., and Diamond, L.S.  (1993) Entamoeba histolytica: a
                    method for isolate identification.  Exp. Parasitol.
                    77:450.). The DNA was mechanically sheared to give a
                    tight size distribution (~2 kb). The v + i method used for
                    the library  construction is described in detail in Smith,
                    H.O. and Venter, J.C.  (Making small insert libraries for
                    whole genome shotgun sequencing  projects. In Genome
                    Sequencing: A Practical Approach, eds. M. Vaudin  and B.
                    Barell, Oxford University Press, 1999)."
BASE COUNT      364 a    107 c    181 g    234 t
ORIGIN

Alignment Scores:
Pred. No.:              6.95e-74      Length:        886
Score:                  726.50        Matches:       149
Percent Similarity:     72.08%        Conservative:  42
Best Local Similarity:  56.23%        Mismatches:    72
Query Match:            17.81%        Indels:        3
DB:                     17            Gaps:          1

US-09-873-409-2 (1-812) x AZ540627 (1-886)

Qy     547 LeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCys 566
           :::::::::   |||:::|||||   ||||||       |||         |||       |||
Db      22 ATTGATAGAATACCAGATATTGATTGTAGATCTACAGCTGGATTAGTTCCAACTGAGTGT 81

Qy     567 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 586
              ||||||:::::      :::|||   |||   ||||||   ||||||:::
Db      82 ATTGGAAATATTAAATTAGAAGATGTTCAATTTAGATATCCAACAAGACCAAATAAACAA 141

Qy     587 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 606
           ||||||   ||||||   |||   |||::::::|||:::|||||||||   ||||||:::
Db     142 ATTCTTGGTGGACTTGACCTTGAAATTAAGAAAGGAGAAACAGTTGCATTAGTAGGAGCA 201

Qy     607 SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGln 626
           ||||||||||||||||||||||||::::::||||||:::||||||:::|||||||||||||
Db     202 TCAGGATGTGGTAAGTCAACTACAATTCAATTAGTCCAAAGAGTCTATGATCCAGTTGGT 261

Qy     627 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 646
           |||:::|||      ||||||   |||   :::|||||||||::::::|||||||||:::
Db     262 GGAAAAGTAACATTAGATGGAAATGATTTAAGAGAGTTGAACTTGAAATGGTTAAGAAAT 321

Qy     647 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 666
           ||||||   :::|||   |||||||||:::||||||   |||:::|||   |||||||||
Db     322 CAAATAGGATTAGTTGGACAAGAACCTATTTTGTTTGCATGTACTATTAGAGAAAATATT 381

Qy     667 AlaTyrGlyAspAsnSerArgValValProLeu---AspGluIleLysGluAlaAlaAsn 685
                 |||               |||      :::|||:::   |||   |||
Db     382 ATGCTTGGAGCTAAAGATGGAGAAACACCAACTGAAGAAGAGATGATTGAATGTGCTAAA 441

Qy     686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 705
              ||||||   |||   ||||||      |||||||||   |||:::|||   ||||||
Db     442 ATGGCAAATGCACATGAATTTATTTCTCATCTTCCAGAAGGATATGATACAATGGTAGGA 501

Qy     706 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725
              :::||||||   ||||||||||||||||||||||||:::||||||||||||||||||
Db     502 GAAAGAGGAGCTGCATTATCAGGAGGACAAAAACAAAGAATTGCTATTGCACGTGCATTG 561

Qy     726 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745
           ::::::||||||   ||||||||||||||||||||||||||||||||||||      |||
Db     562 ATTAGAAAACCTACAATTTTATTACTTGATGAAGCTACATCAGCACTTGATACACAAAGT 621

Qy     746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765
           ||||||:::||||||   ||||||:::||||||      |||||||||   ::::::|||
Db     622 GAAAAGATTGTACAACAAGCACTTGAAAAAGCTAGTCAAGGAAGAACAACAATTATTGTA 681

Qy     766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeu-IleValValLeuHisAsnGlyLy 785
              |||||||||:::   ::::::||||||:::   |||   |||   |||   |||::
Db     682 GCACATAGATTAACAACTGTTAGAAATGCAAATAGAAATTTGTGTATTCCATCAAGGAGA 741

Qy     785 sIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLe 805
            :|||   |||||||||||||||||||||||||:::               |||:::   ||
Db     742 AATTATTGAACAAGGAACACATCAAGAATTAATGGATTT-GAAAGCACATATTATGGATT 800

Qy     805 uValAsnAlaGln 809
            ||||     |||
Db     801 AGTTAAAAGACAA 813

RESULT 12
AZ687805/c
LOCUS       AZ687805              880 bp    DNA     linear   GSS 14-DEC-2000
DEFINITION  ENTLU52TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   AZ687805
VERSION     AZ687805.1  GI:11824951
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
  ORGANISM  Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 880)
  AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Forward
            Class: shotgun
            High quality sequence start: 30
            High quality sequence stop: 858.
FEATURES             Location/Qualifiers
     source          1..880
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a  method described by Clark and Diamond (Clark,
                     C.G., and Diamond, L.S.  (1993) Entamoeba histolytica: a
                     method for isolate identification.  Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + i method used for
                     the library  construction is described in detail in Smith,
                     H.O. and Venter, J.C.  (Making small insert libraries for
                     whole genome shotgun sequencing  projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin  and B.
                     Barell, Oxford University Press, 1999)."
BASE COUNT      247 a    167 c    122 g    344 t
ORIGIN

Alignment Scores:
Pred. No.:              1.54e-73      Length:        880
Score:                  723.50        Matches:       151
Percent Similarity:     69.93%        Conservative:  42
Best Local Similarity:  54.71%        Mismatches:    82
Query Match:            17.74%        Indels:        2
DB:                     17            Gaps:          1

US-09-873-409-2 (1-812) x AZ687805 (1-880)
```

```
Qy     536 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSer 555
           ||||||:::   |||      ::::::       :::::::::   |||:::||||||
Db     831 GCTAAAGCTTCTGCATATAGAATTTATCAAACAATTGATAGAATCCCAGATATTGATTGT 772

Qy     556 ArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 575
           ||||||      |||:::   |||         |||   ||||||:::          :::|||
Db     771 AGATCTACAGCTGGTGAATGTCCAACTGAGTGTAATGGAAATATTACATTAGAAGATGTT 712

Qy     576 SerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIle 595
              |||   ||||||   ||||||            ||||||   ||||||   |||   |||
Db     711 CAATTTAGATATCCAACAAGACCAACTAAACAAATTCTTGGTGGACTTGATCTTGAAATT 652

Qy     596 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 615
           ::::::|||:::|||||||||      ||||||:::||||||||||||||||||||||||::::::
Db     651 AAGAAAGGACAAACAGTTGCATTAGTAGGAGCATCAGGATGTGGTAAATCAACTACTATT 592

Qy     616 GlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAsp 635
           ||||||:::||||||   ||||||||||||   |||   |||      ||||||   |||
Db     591 CAATTAGTCCAAAGAAATTATGATCCAGTTGGTGGGTCAGTCAAATTAGATGGAAAAGAT 532

Qy     636 AlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 655
           ::::::|||||||::::::|||||||||:::||||||      :::|||   |||||||||
Db     531 TTAAGAGATTTAAATATCAAATGGTTAAGAAATCAAATAGGATTAGTTGGACAAGAACCT 472

Qy     656 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValVal 675
           :::||||||   |||:::|||   |||||||||         |||
Db     471 ATTTTGTTTGCATGTACTATTAGAGAAAATATTATGCTTGGAGCTAGAGATGGAGAAACA 412

Qy     676 ProLeu---AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 694
           |||      :::|||:::   |||   |||      ||||||   |||   ||||||
Db     411 CCAACTGAAGAAGAGATGATTGAATGTGCTAAAATGGCAAATGCACATGAATTTATTTCT 352

Qy     695 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 714
              |||||||||   |||:::|||   ||||||   |||||||||   ||||||||||||
Db     351 CATCTTCCAGAAGGATATGATACAATGGTAGGAGAAAAAGGAGCTGCATTATCAGGAGGA 292

Qy     715 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 734
           |||||||||||:::|||||||||||||||||||:::::|||||   ||||||||||||
Db     291 CAAAAACAAAGAATTGCTATTGCACGTGCATTGATTAGAAAACCTACAATTTTATTACTT 232

Qy     735 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 754
           |||||||||||||||||||||||         |||||||||:::||||||   ||||||:::
Db     231 GATGAAGCTACATCAGCACTTGATACACAAAGTGAAAAGATTGTACAACAAGCACTTGAA 172

Qy     755 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 774
           ||||||      |||||||||   :::||||||   |||||||||:::     ::::::|||
Db     171 AAAGCTAGTCAAGGAAGAACAACAATTGTTGTAGCACATAGATTAACAACTGTTAGAAAT 112

Qy     775 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 794
           |||      |||   |||   |||   |||:::|||         ||||||||||||||||
Db     111 GCAAGTAGAATTTGTGTATTCCATCAAGGAGAAATTATTAA-CAAGGAACACATCAAGAA 53

Qy     795 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 810
           |||            :::|||   :::   :::      ||||||:::
Db      52 TTAAGGATTTAAAGCAACATATTATGGATTATTTTTCCAGCACAAACT 5
```

```
RESULT 13
BQ720763
LOCUS       BQ720763                 939 bp    mRNA    linear   EST 16-JUL-2002
DEFINITION  AGENCOURT_8233170 Lupski_dorsal_root_ganglion Homo sapiens cDNA
            clone IMAGE:6184557 5', mRNA sequence.
ACCESSION   BQ720763
VERSION     BQ720763.1  GI:21859660
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 939)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13573  row: o  column: 22
            High quality sequence stop: 646.
FEATURES             Location/Qualifiers
     source          1..939
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6184557"
                     /clone_lib="Lupski_dorsal_root_ganglion"
                     /sex="male"
                     /tissue_type="dorsal root ganglia"
                     /dev_stage="adult, 36 yr"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.7 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
BASE COUNT      272 a    202 c    221 g    244 t
ORIGIN

Alignment Scores:
Pred. No.:              2.95e-73        Length:          939
Score:                  721.50          Matches:         140
Percent Similarity:     74.44%          Conservative:    61
Best Local Similarity:  51.85%          Mismatches:      66
Query Match:            17.69%          Indels:          3
DB:                     14              Gaps:            1

US-09-873-409-2 (1-812) x BQ720763 (1-939)

Qy     314 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 333
           :::||||||||||||||||||   |||||||||::::::   |||:::::::||||||
Db       3 AAAGCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGA 62

Qy     334 GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 353
           ||||||::::::||||||||||||   :::|||:::||||||   |||||||||   |||
Db      63 CAGGATGTGAGTTGGTTTGATGAGCCTAAAAACACCACTGGAGCATTGACTACCAGGCTC 122

Qy     354 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 373
           |||   |||   ||||||::::::||||||   |||||||||:::   |||:::||||||
Db     123 GCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAG 182

Qy     374 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr 393
           |||      |||:::|||      :::||||||||||||||||||||||||||::::::|||
Db     183 AATATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACA 242

Qy     394 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 413
              |||:::|||:::|||   |||::::::|||:::   |||::::::|||
Db     243 CTGTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATG 302

Qy     414 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 433
           ::::::|||   |||   |||||||||:::||||||:::   :::||||||||||||||||
Db     303 TTGTCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGTTCTGGGAAGATCGCTACT 362

Qy     434 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 453
           ||||||:::||||||   ||||||:::||||||||||||:::|||:::   ||||||
Db     363 GAAGCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACAT 422

Qy     454 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 473
           ||||||   :::   ||||||         :::||||||:::   :::||||||   |||
```

```
Db     423 ATGTATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTT 482

Qy     474 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 493

Db     483 GGAATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGG 542

Qy     494 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 513

Db     543 TTTGGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTT 602

Qy     514 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 533

Db     603 TCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTAT 662

Qy     534 Ser-LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIl 553

Db     663 GCCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGAT 722

Qy     553 eAspSerArgSer-GlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGlu---P 572

Db     723 TGACAGCTACAGCACCGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACCATTT 782

Qy     572 heArgGluValSerPhePheTyrPro 580

Db     783 GGTGAAAGTTGTATTTCAACTATCCC 808
```

```
RESULT 14
AZ541090
LOCUS       AZ541090                897 bp    DNA     linear   GSS 14-NOV-2000
DEFINITION  ENTDS67TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   AZ541090
VERSION     AZ541090.1  GI:11148493
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
  ORGANISM  Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 897)
  AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Reverse
            Class: shotgun
            High quality sequence start: 81
            High quality sequence stop: 890.
FEATURES             Location/Qualifiers
     source          1..897
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a  method described by Clark and Diamond (Clark,
                     C.G., and Diamond, L.S.  (1993) Entamoeba histolytica: a
                     method for isolate identification.  Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + i method used for
                     the library  construction is described in detail in Smith,
                     H.O. and Venter, J.C.  (Making small insert libraries for
                     whole genome shotgun sequencing  projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin  and B.
                     Barell, Oxford University Press, 1999)."
BASE COUNT      335 a    110 c    177 g    275 t
ORIGIN

Alignment Scores:
Pred. No.:                 3.58e-73      Length:         897
Score:                     720.50        Matches:        143
Percent Similarity:        74.59%        Conservative:   39
Best Local Similarity:     58.61%        Mismatches:     61
Query Match:               17.66%        Indels:         1
DB:                        17            Gaps:           1

US-09-873-409-2 (1-812) x AZ541090 (1-897)
```

```
Qy     567 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 586

Db      94 AAAGGTGAAATTGAATTTAAAGACATTTGTTTCAGATATCCAACAAGACCAGACAATTCT 153

Qy     587 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 606

Db     154 GTCTTGAAAGGTATTTCATTCAAAGTAGAACAAGGAAAAACTGTTGCATTAGTAGGAGCA 213

Qy     607 SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGln 626

Db     214 TCAGGATGTGGTAAATCAACATCAGTTCAGTTGATTGAAAGATTTTATGATCCAACACAT 273

Qy     627 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 646

Db     274 GGAGATGTATTATTAGATGGACATAATATCAAAGATTTGAATATTCATTTCTTAAGAAGT 333

Qy     647 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 666

Db     334 CAAATTGGAATGGTAGGACAAGAACCAGTATTATTTGCTGAAAGTGTTATGGATAATATT 393

Qy     667 AlaTyrGlyAspAsnSerArgVal---ValProLeuAspGluIleLysGluAlaAlaAsn 685

Db     394 AGAAGAGGAGTACCTAAAGGAGTTGAAGTAAGTAATGAACAAATTTATGCTGCTGCTAAA 453

Qy     686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 705

Db     454 ATGGCAAATGCACATGACTTTATTTCAGCAATGCCAGAAGGATATAACACAATGGTAGGT 513

Qy     706 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725

Db     514 GATAGAGGTGCACAAATTTCAGGAGGACAAAAACAAAGAATTGCTATTGCACGTGCATTG 573

Qy     726 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745

Db     574 ATTAGAAATCCAAAAGTGTTATTACTCGATGAAGCTACATCAGCACTTGATTCAGAAAGT 633

Qy     746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765

Db     634 GAAAAGATTGTACAAGATGCACTTGACAAAGCAGCAAAAGGAAGAACAACAATTGTAATT 693

Qy     766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785

Db     694 GCACATAGATTATCAACTATTCAAAATGCAGATCAAATATGTGTTATTATGAGAGGAAGA 753

Qy     786 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 805

Db     754 ATTGCAGAAAGAGGAACACATCAAGAGTTATTAGATTTGAAAGGATTTTATTATACACTT 813

Qy     806 ValAsnAlaGln 809

Db     814 GCTATGCAACAA 825
```

```
RESULT 15
BH155700
LOCUS       BH155700                913 bp    DNA     linear   GSS 24-SEP-2001
DEFINITION  ENTRO54TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   BH155700
VERSION     BH155700.1  GI:15727822
```

```
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
  ORGANISM  Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 913)
  AUTHORS   Loftus,B., Wang,Z., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library (2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Reverse
            Class: shotgun
            High quality sequence start: 7
            High quality sequence stop: 805.
FEATURES             Location/Qualifiers
     source          1..913
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a  method described by Clark and Diamond (Clark,
                     C.G., and Diamond, L.S.  (1993) Entamoeba histolytica: a
                     method for isolate identification.  Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + i method used for
                     the library  construction is described in detail in Smith,
                     H.O. and Venter, J.C.  (Making small insert libraries for
                     whole genome shotgun sequencing  projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin  and B.
                     Barell, Oxford University Press, 1999)."
BASE COUNT      359 a    118 c    185 g    251 t
ORIGIN

Alignment Scores:
Pred. No.:              4.17e-72       Length:          913
Score:                  711.50         Matches:         146
Percent Similarity:     69.37%         Conservative:    42
Best Local Similarity:  53.87%         Mismatches:      82
Query Match:            17.44%         Indels:          2
DB:                     17             Gaps:            1

US-09-873-409-2 (1-812) x BH155700 (1-913)

Qy     533 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 552
           ::::::   ||||||:::   |||      ::::::      :::::::::   |||:::
Db      29 TTTGCTACAGCTAAAGCTTCTGCATATAGAATTTATCAAACAATTGATAGAATCCCAGAT 88

Qy     553 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 572
           ||||||   ||||||         |||:::   |||         |||   ||||||:::
Db      89 ATTGATTGTAGATCTACAGCTGGTGAATGTCCAACTGAGTGTAATGGAAATATTACATTA 148

Qy     573 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 592
              :::|||   |||   ||||||   ||||||         ||||||   ||||||
Db     149 GAAGATGTTCAATTTAGATATCCAACAAGACCAACTAAACAAATTCTTGGTGGACTTGAT 208

Qy     593 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 612
           |||   |||::::::|||:::|||||||||   ||||||:::||||||||||||||||||
Db     209 CTTGAAATTAAGAAAGGACAAACAGTTGCATTAGTAGGAGCATCAGGATGTGGTAAATCA 268

Qy     613 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 632
           |||::::::||||||:::||||||   ||||||||||||   |||   |||      |||
Db     269 ACTACTATTCAATTAGTCCAAAGAAATTATGATCCAGTTGGTGGGTCAGTCAAATTAGAT 328

Qy     633 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 652
           |||   |||      ::::::||||||::::::|||||||||:::||||||      :::|||
Db     329 GGAAAAGATTTAAGAGATTTAAATATCAAATGGTTAAGAAATCAAATAGGATTAGTTGGA 388

Qy     653 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 672
           |||||||||:::||||||      |||:::|||   |||||||||         |||
Db     389 CAAGAACCTATTTTGTTTGCATGTACTATTAGAGAAAATATTATGCTTGGAGCTAGAGAT 448

Qy     673 ArgValValProLeu---AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSer 691
                    |||      :::|||:::   |||   |||      ||||||   |||
Db     449 GGAGAAACACCAACTGAAGAAGAGATGATTGAATGTGCTAAAATGGCAAATGCACATGAA 508

Qy     692 PheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeu 711
           ||||||         |||||||||   |||:::|||   ||||||   |||||||||   |||
Db     509 TTTATTTCTCATCTTCCAGAAGGATATGATACAATGGTAGGAGAAAAAGGAGCTGCATTA 568

Qy     712 SerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIle 731
           |||||||||||||||||||:::||||||||||||||||||::::::||||||   |||
Db     569 TCAGGAGGACAAAAACAAAGAATTGCTATTGCACGTGCATTGATTAGAAAACCTACAATT 628

Qy     732 LeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHis 751
           ||||||||||||||||||||||||||||||         |||||||||:::||||||
Db     629 TTATTACTTGATGAAGCTACATCAGCACTTGATACACAAAGTGAAAAGATTGTACAACAA 688

Qy     752 AlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAla 771
           ||||||:::||||||      |||||||||      :::||||||   |||||||||:::
Db     689 GCACTTGAAAAAGCTAGTCAAGGAAGAACAACAATTGTTGTAGCACATAGATTAACAACT 748

Qy     772 IleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThr 791
           ::::::||||||            |||   |||   |||:::|||         ||||||
Db     749 GTTAGAAATGCAAGTAGAAATTGTGTATTCCATCAAGGAGAAATTAT-TTGACAGGAACA 807

Qy     792 HisGlnGluLeuLeuArgAsnArgAspIleTyr 802
           |||   ||||||:::      :::   |||
Db     808 CATTCAGAATTAATGGATTTGAAGGGGACATAT 840


Search completed: March 31, 2003, 13:49:57
Job time : 5612.21 secs
```

GenCore version 5.1.4_p5_4578


```
OM protein - nucleic search, using frame_plus_p2n model

Run on:          March 31, 2003, 02:24:20 ; Search time 118.913 Seconds
                                           (without alignments)
                                           2094.145 Million cell updates/sec

Title:           US-09-873-409-2
Perfect score:   4079
Sequence:        1 MVDENDIRALNVRHYRDHIG..........QELLRNRDIYFKLVNAQSVQ 812

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters:      882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115422_19259/app_query.fasta_1.7544
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_578_@runat_27032003_115422_19259 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       Issued_Patents_NA:*
                 1:  /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
                 2:  /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
                 3:  /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
                 4:  /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
                 5:  /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
                 6:  /cgn2_6/ptodata/2/ina/backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2436	59.7	4264	2	US-08-784-649A-1	Sequence 1, Appli
2	2436	59.7	4264	2	US-08-784-649A-5	Sequence 5, Appli
3	2436	59.7	4646	1	US-08-181-471-2	Sequence 2, Appli
4	2436	59.7	4669	6	5206352-3	Patent No. 5206352
5	2436	59.7	6505	2	US-08-793-610-5	Sequence 5, Appli
6	2436	59.7	9318	2	US-08-793-610-6	Sequence 6, Appli
7	2413.5	59.2	4233	3	US-09-120-513-1	Sequence 1, Appli
8	2413.5	59.2	4233	4	US-09-450-105-1	Sequence 1, Appli
9	2412	59.1	4669	2	US-08-752-447-1	Sequence 1, Appli
10	2412	59.1	4669	4	US-09-316-167-1	Sequence 1, Appli
11	2397	58.8	4669	2	US-08-583-276-18	Sequence 18, Appl
12	2280	55.9	2726	1	US-08-461-823-1	Sequence 1, Appli
13	1690.5	41.4	4002	2	US-08-996-545-1	Sequence 1, Appli
14	1690.5	41.4	4002	2	US-08-996-545-3	Sequence 3, Appli
15	1690.5	41.4	4002	4	US-09-328-320-1	Sequence 1, Appli
16	1690.5	41.4	4002	4	US-09-328-320-3	Sequence 3, Appli
17	1661	40.7	4047	2	US-08-612-734B-1	Sequence 1, Appli
18	1618	39.7	4800	2	US-08-612-734B-3	Sequence 3, Appli
19	1570.5	38.5	3924	1	US-08-395-246C-1	Sequence 1, Appli
20	1474	36.1	4224	1	US-08-612-521-1	Sequence 1, Appli
21	1255.5	30.8	3909	1	US-08-232-537-1	Sequence 1, Appli
22	1223.5	30.0	3924	2	US-08-996-644-3	Sequence 3, Appli
23	1223.5	30.0	3924	3	US-09-352-552-3	Sequence 3, Appli
24	1223.5	30.0	3927	2	US-08-996-644-1	Sequence 1, Appli
25	1223.5	30.0	3927	3	US-09-352-552-1	Sequence 1, Appli
26	1172	28.7	6143	1	US-08-612-521-3	Sequence 3, Appli
27	1116.5	27.4	3792	4	US-09-351-224B-10	Sequence 10, Appl
28	1032.5	25.3	3999	4	US-09-351-224B-9	Sequence 9, Appli
29	915	22.4	13188	4	US-08-961-527-70	Sequence 70, Appl
30	824	20.2	6492	4	US-08-961-527-188	Sequence 188, App
31	793	19.4	7760	4	US-08-961-527-63	Sequence 63, Appl
32	776	19.0	7186	4	US-08-961-527-39	Sequence 39, Appl
33	763	18.7	4403765	4	US-09-103-840A-2	Sequence 2, Appli
34	763	18.7	4411529	4	US-09-103-840A-1	Sequence 1, Appli
35	678	16.6	2376	1	US-08-394-880B-1	Sequence 1, Appli
36	644	15.8	2244	4	US-09-061-764A-18	Sequence 18, Appl
37	623.5	15.3	2061	4	US-09-061-764A-17	Sequence 17, Appl
38	608.5	14.9	4781	2	US-09-001-273-1	Sequence 1, Appli
39	608.5	14.9	4781	4	US-08-843-459A-1	Sequence 1, Appli
40	608.5	14.9	4847	3	US-09-061-400-1	Sequence 1, Appli
41	608.5	14.9	5120	3	US-08-772-270A-6	Sequence 6, Appli
42	608.5	14.9	8370	2	US-08-488-706-1	Sequence 1, Appli
43	601	14.7	1749	4	US-09-134-001C-1893	Sequence 1893, Ap
44	600.5	14.7	1743	4	US-09-134-001C-2774	Sequence 2774, Ap
45	594.5	14.6	7721	3	US-08-772-270A-14	Sequence 14, Appl

ALIGNMENTS

```
RESULT 1
US-08-784-649A-1
; Sequence 1, Application US/08784649A
; Patent No. 5830697
;  GENERAL INFORMATION:
;    APPLICANT:  Sikic, Branimir I
;    APPLICANT:  Chen, Gang
;    TITLE OF INVENTION:  P-GLYCOPROTEIN MUTANT RESISTANT TO
;    TITLE OF INVENTION:  CYCLOSPORIN MODULATION
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  2200 Sand Hill Road
;      CITY:  Menlo Park
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/784,649A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sherwood, Pamela J
;      REGISTRATION NUMBER:  Reg.No. 5830697 36,677
;      REFERENCE/DOCKET NUMBER:  06037/007001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-322-5070
;      TELEFAX:  415-854-0875
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4264 base pairs
```

```
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-784-649A-1

Alignment Scores:
Pred. No.:                6.88e-275     Length:          4264
Score:                    2436.00       Matches:         466
Percent Similarity:       76.79%        Conservative:    166
Best Local Similarity:    56.62%        Mismatches:      177
Query Match:              59.72%        Indels:          14
DB:                       2             Gaps:            4

US-09-873-409-2 (1-812) x US-08-784-649A-1 (1-4264)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db   1495 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 1554

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db   1555 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 1614

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db   1615 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1674

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db   1675 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 1734

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db   1735 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 1794

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db   1795 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 1854

Qy    122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db   1855 CTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGTTTGTCTACAGTT 1914

Qy    142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db   1915 CGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 1974

Qy    162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178
Db   1975 GATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATGCAGACAGCAGGA 2034

Qy    179 ---------AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
Db   2035 AATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATTGATGCCTTGGAA 2094

Qy    196 ArgLysThrAsnSerLeuProLeuHisSerVal-----------------LysSerIle 209
Db   2095 ATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGTC 2154

Qy    210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228
Db   2155 CGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTGGATGAAAGTATA 2214

Qy    229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
Db   2215 CCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCCTTATTTTGTT 2274

Qy    249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db   2275 GTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTTGCAATAATATTT 2334

Qy    269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
Db   2335 TCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAACGACAGAATAGT 2394

Qy    288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db   2395 AACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATTACATTTTTCCTT 2454

Qy    308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db   2455 CAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTT 2514

Qy    328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
Db   2515 TTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAAAACACCACTGGA 2574

Qy    348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db   2575 GCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGG 2634

Qy    368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db   2635 CTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATC 2694

Qy    388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db   2695 TATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGA 2754

Qy    408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db   2755 GTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGT 2814

Qy    428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db   2815 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAG 2874

Qy    448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db   2875 GAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 2934

Qy    468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db   2935 AGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCC 2994

Qy    488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db   2995 TATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAG 3054

Qy    508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db   3055 GATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGT 3114

Qy    528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db   3115 TCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATT 3174

Qy    548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db   3175 GAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAA 3234

Qy    568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db   3235 GGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTG 3294

Qy    588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db   3295 CTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGT 3354

Qy    608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db   3355 GGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGG 3414

Qy    628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
```

```
Db    3415 AAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACAC 3474

Qy     648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
           :::   ||||||   |||||||||:::||||||:::||||||||||||||||||||||
Db    3475 CTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCC 3534

Qy     668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
           ||||||||||||||||||||||||            :::||||||         ||||||         |||
Db    3535 TATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATTGTGAGGGCAGCAAAGGAGGCC 3594

Qy     688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
           |||||||||:::|||||||||   ||||||   ||||||:::|||:::||||||      |||
Db    3595 AACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAA 3654

Qy     708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
           |||   |||||||||||||||||||||||||||:::||||||||||||||||||:::::::
Db    3655 GGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGA 3714

Qy     728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
           :::|||   |||||||||||||||||||||||||||||||||||||      :::|||||||||
Db    3715 CAGCCTCATATTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAG 3774

Qy     748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
           |||||||||   ||||||||||||||||||   |||||||||||:::|||:::      |||
Db    3775 GTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCAC 3834

Qy     768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
           |||||||||   |||||||||||||||||||||||||||||||      ||||||:::::::|||
Db    3835 CGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGCAGAGTCAAG 3894

Qy     788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
           |||   ||||||||||||:::||||||         :::   |||||||||      :::|||:::
Db    3895 GAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGT 3954

Qy     808 AlaGlnSer 810
              |||:::
Db    3955 GTCCAGGCT 3963
```

RESULT 2
US-08-784-649A-5
; Sequence 5, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
;   APPLICANT: Sikic, Branimir I
;   APPLICANT: Chen, Gang
;   TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
;   TITLE OF INVENTION: CYCLOSPORIN MODULATION
;   NUMBER OF SEQUENCES: 5
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Fish & Richardson
;   STREET: 2200 Sand Hill Road
;   CITY: Menlo Park
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94025
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/784,649A
;   FILING DATE:
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Sherwood, Pamela J
;   REGISTRATION NUMBER: Reg.No. 5830697 36,677
;   REFERENCE/DOCKET NUMBER: 06037/007001
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415-322-5070
;   TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 5:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4264 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
US-08-784-649A-5

```
Alignment Scores:
Pred. No.:              6.88e-275      Length:         4264
Score:                  2436.00        Matches:        466
Percent Similarity:     76.79%         Conservative:   166
Best Local Similarity:  56.62%         Mismatches:     177
Query Match:            59.72%         Indels:         14
DB:                     2              Gaps:           4

US-09-873-409-2 (1-812) x US-08-784-649A-5 (1-4264)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
           ||||||      ||||||||||      :::|||||||||         |||:::   |||||||||
Db    1495 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 1554

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           ||||||||||||||||||||||||   |||||||||:::      ||||||:::|||||||||
Db    1555 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 1614

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
           ::::::||||||   :::|||:::|||:::|||      :::|||||||||||||||||||||
Db    1615 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1674

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           ||||||:::   |||:::||||||:::||||||||||||||||:::||||||||||:::|||
Db    1675 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 1734

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1735 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 1794

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
           :::||||||||||||||||||||||||||:::||||||::::::   ||||||   |||
Db    1795 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 1854

Qy     122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
           |||:::||||||   |||||||||||||||||||:::||||||||||||||||||:::
Db    1855 CTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGTTTGTCTACAGTT 1914

Qy     142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
           |||:::||||||:::|||         ||||||::::::   |||||||||   |||
Db    1915 CGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 1974

Qy     162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178
              |||||||||   ::::::|||:::|||:::   ||||||      |||
Db    1975 GATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATGCAGACAGCAGGA 2034

Qy     179 ---------AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
                    :::::::::   |||   ::::::   :::|||         :::   |||
Db    2035 AATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATTGATGCCTTGGAA 2094

Qy     196 ArgLysThrAsnSerLeuProLeuHisSerVal------------------LysSerIle 209
                 :::|||            :::                     :::|||:::
Db    2095 ATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGTC 2154

Qy     210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228
           :::   |||      |||:::         :::   ::::::   :::   |||:::
Db    2155 CGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTGGATGAAAGTATA 2214

Qy     229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
           |||   ||||||         :::|||:::|||||||||      |||||||||:::   |||
Db    2215 CCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCCTTATTTTGTT 2274

Qy     249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
           :::|||         :::::::::||||||   :::   |||   |||:::|||||||||
Db    2275 GTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTTGCAATAATATTT 2334
```

```
Qy     269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
           :::||||||||   :::|||          :::|||       |||   :::   ::::::
Db    2335 TCAAAGATTATAGGGGTTTTACAAGAATTGATGATCCTGAAACAAAACGACAGAATAGT 2394

Qy     288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
              :::::|||:::::|||:::    ||||||:::|||    |||:::::::::|||:::
Db    2395 AACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATTACATTTTTCCTT 2454

Qy     308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
           ||||||      :::|||:::||||||||||||||||||    |||||||||::::::
Db    2455 CAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTT 2514

Qy     328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
           |||::::::||||||   ||||||:::::::|||||||||:::    :::|||:::||||||
Db    2515 TTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAAAACACCACTGGA 2574

Qy     348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
              |||||||||   ||||||   |||   ||||||::::::||||||    |||||||||
Db    2575 GCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGG 2634

Qy     368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
           :::    |||:::|||||||||||      |||:::|||        :::|||||||||||||||
Db    2635 CTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATC 2694

Qy     388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
           |||||||||::::::|||   |||:::|||:::|||    |||::::::|||:::    |||
Db    2695 TATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGA 2754

Qy     408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
           ::::::|||         ::::::|||   |||   |||||||||:::||||||:::
Db    2755 GTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGT 2814

Qy     428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
           |||||||||||||||||||||||:::||||||   ||||||:::|||||||||||:::
Db    2815 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAG 2874

Qy     448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
           |||:::   ||||||   ||||||   :::   ||||||         :::||||||:::
Db    2875 GAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 2934

Qy     468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
           :::||||||   |||   |||      :::::::|||:::   |||   :::||||||:::
Db    2935 AGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCC 2994

Qy     488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
           ||||||      |||||||||||||||||||||:::            |||:::   |||
Db    2995 TATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAG 3054

Qy     508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
              :::   :::||||||:::|||:::   :::|||||||||||:::|||:::
Db    3055 GATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGT 3114

Qy     528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
                 ||||||:::|||:::||||||||     |||||||||:::       ::::::
Db    3115 TCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATT 3174

Qy     548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
           ||||||   |||   |||||||||   |||   ||||||       |||:::|||   |||
Db    3175 GAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAA 3234

Qy     568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
           ||||||:::   |||   ||||||   |||   ||||||   ||||||||||:::   :::
Db    3235 GGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTG 3294

Qy     588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
           |||:::||||||||||||   :::::::::|||:::|||:::|||   ||||||||||||
Db    3295 CTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGT 3354

Qy     608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
           ||||||||||||||||||   ||||||||||||:::|||   ||||||||||:::   |||
Db    3355 GGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGG 3414

Qy     628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
           :::||||||   ||||||   :::   |||   |||||||||||||||||||:::
Db    3415 AAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACAC 3474

Qy     648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
           :::   ||||||   ||||||||||:::||||||:::|||||||||||||||||||||
Db    3475 CTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCC 3534

Qy     668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
           ||||||||||||||||||||||||       :::||||||      ||||||      |||
Db    3535 TATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATTGTGAGGGCAGCAAAGGAGGCC 3594

Qy     688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
           |||||||||:::|||||||||   ||||||   ||||||:::|||:::||||||   |||
Db    3595 AACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAA 3654

Qy     708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
           |||   |||||||||||||||||||||||||:::|||||||||||||||||||::::::
Db    3655 GGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGA 3714

Qy     728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
           :::|||   |||||||||||||||||||||||||||||||||      :::|||||||||
Db    3715 CAGCCTCATATTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAG 3774

Qy     748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
           |||||||||   ||||||||||||||||||   ||||||||||||:::|||:::   |||
Db    3775 GTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCAC 3834

Qy     768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
           |||||||||   |||||||||||||||||||||||||||      ||||||::::::|||
Db    3835 CGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGCAGAGTCAAG 3894

Qy     788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
           |||   |||||||||||:::||||||      :::   |||||||||   :::|||:::
Db    3895 GAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGT 3954

Qy     808 AlaGlnSer 810
              |||:::
Db    3955 GTCCAGGCT 3963
```

RESULT 3
US-08-181-471-2
; Sequence 2, Application US/08181471
; Patent No. 5641508
; GENERAL INFORMATION:
;   APPLICANT: Li, Lingna
;   APPLICANT: Lishko, Valeryi K.
;   TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL
;   TITLE OF INVENTION: COMPOSITIONS TO HAIR FOLLICLES
;   NUMBER OF SEQUENCES: 3
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Thomas Fitting
;   STREET: 12526 High Bluff Drive, Suite 300
;   CITY: San Diego
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 92130
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/181,471
;   FILING DATE: 13-JAN-1994
;   CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/041,553
;   FILING DATE: 02-APR-1993
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Fitting, Thomas
;   REGISTRATION NUMBER: 34,163
;   REFERENCE/DOCKET NUMBER: ANT0029P
;   TELECOMMUNICATION INFORMATION:

```
;      TELEPHONE:  619-792-3680
;      TELEFAX:  619-792-8477
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  4646 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  425..4267
US-08-181-471-2

Alignment Scores:
Pred. No.:               8e-275      Length:         4646
Score:                   2436.00     Matches:        466
Percent Similarity:      76.79%      Conservative:   166
Best Local Similarity:   56.62%      Mismatches:     177
Query Match:             59.72%      Indels:         14
DB:                      1           Gaps:           4

US-09-873-409-2 (1-812) x US-08-181-471-2 (1-4646)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db   1781 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 1840

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db   1841 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 1900

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db   1901 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1960

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db   1961 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 2020

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db   2021 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 2080

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db   2081 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 2140

Qy    122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db   2141 CTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGTTTGTCTACAGTT 2200

Qy    142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db   2201 CGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 2260

Qy    162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178
Db   2261 GATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATGCAGACAGCAGGA 2320

Qy    179 ---------AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
Db   2321 AATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATTGATGCCTTGGAA 2380

Qy    196 ArgLysThrAsnSerLeuProLeuHisSerVal------------------LysSerIle 209
Db   2381 ATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGTC 2440

Qy    210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228
Db   2441 CGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTGGATGAAAGTATA 2500

Qy    229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
Db   2501 CCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCCTTATTTTGTT 2560

Qy    249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db   2561 GTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTTGCAATAATATTT 2620

Qy    269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
Db   2621 TCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAACGACAGAATAGT 2680

Qy    288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db   2681 AACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATTACATTTTTCCTT 2740

Qy    308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db   2741 CAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTT 2800

Qy    328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
Db   2801 TTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAAAACACCACTGGA 2860

Qy    348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db   2861 GCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGG 2920

Qy    368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db   2921 CTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATC 2980

Qy    388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db   2981 TATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGA 3040

Qy    408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db   3041 GTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGT 3100

Qy    428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db   3101 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAG 3160

Qy    448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db   3161 GAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 3220

Qy    468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db   3221 AGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCC 3280

Qy    488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db   3281 TATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAG 3340

Qy    508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db   3341 GATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGT 3400

Qy    528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db   3401 TCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATT 3460

Qy    548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db   3461 GAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAA 3520

Qy    568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db   3521 GGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTG 3580
```

```
Qy     588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
           |||:::||||||||||||   :::::::::|||:::|||:::|||   ||||||||||||
Db    3581 CTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGT 3640

Qy     608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
           ||||||||||||||||||   ||||||||||||:::|||   |||||||||:::   |||
Db    3641 GGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGG 3700

Qy     628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
           :::||||||   ||||||   :::   |||   |||||||||||||||||||||:::
Db    3701 AAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACAC 3760

Qy     648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
           :::   ||||||   |||||||||:::||||||:::||||||||||||||||||||||||
Db    3761 CTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCC 3820

Qy     668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
           ||||||||||||||||||||||||         :::||||||      ||||||   |||
Db    3821 TATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCC 3880

Qy     688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
           |||||||||:::|||||||||   ||||||   |||||||||:::||||||   |||
Db    3881 AACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAA 3940

Qy     708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
           |||   |||||||||||||||||||||||||||:::|||||||||||||||||:::::::
Db    3941 GGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGA 4000

Qy     728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
           :::|||   ||||||||||||||||||||||||||||||||||||   :::|||||||||
Db    4001 CAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAG 4060

Qy     748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
           |||||||||   ||||||||||||||||||   ||||||||||||:::|||:::   |||
Db    4061 GTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCAC 4120

Qy     768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
           |||||||||   ||||||||||||||||||||||||||||||   ||||||:::::::|||
Db    4121 CGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGCAGAGTCAAG 4180

Qy     788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
           |||   ||||||||||||:::||||||         :::   |||||||||   :::|||:::
Db    4181 GAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGT 4240

Qy     808 AlaGlnSer 810
              |||:::
Db    4241 GTCCAGGCT 4249
```

RESULT 4
5206352-3
;Patent No. 5206352
;   APPLICANT: Roninson, Igor B.;Pastan Ira H.;Gottesman,
;Michael M.
;   TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA
;SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
;   NUMBER OF SEQUENCES: 4
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/622,836
;   FILING DATE: 24-SEP-1990
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 892,575
;   FILING DATE: 01-AUG-1986
;   APPLICATION NUMBER: 845,610
;   FILING DATE: 28-MAR-1986
;SEQ ID NO:3:
;   LENGTH: 4669
5206352-3

Alignment Scores:

Pred. No.:	8.07e-275	Length:	4669
Score:	2436.00	Matches:	466
Percent Similarity:	76.79%	Conservative:	166
Best Local Similarity:	56.62%	Mismatches:	177
Query Match:	59.72%	Indels:	14
DB:	6	Gaps:	4

US-09-873-409-2 (1-812) x 5206352-3 (1-4669)

```
Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
           ||||||      ||||||||||||   :::||||||||||||      |||:::   ||||||||
Db    1781 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 1840

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           ||||||||||||||||||||||||   |||||||||:::   ||||||:::||||||||
Db    1841 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 1900

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
           ::::::||||||   :::|||:::|||:::|||   :::||||||||||||||||||||||
Db    1901 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1960

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           ||||||:::   |||:::||||||||||||||||||:::||||||||||||:::|||
Db    1961 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 2020

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   .2021 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 2080

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
           :::||||||||||||||||||||||||||||:::||||||::::::   ||||||   |||
Db    2081 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 2140

Qy     122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
           |||:::||||||   ||||||||||||||||||||||:::|||||||||||||||||:::
Db    2141 CTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGTTTGTCTACAGTT 2200

Qy     142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
           |||:::||||||:::|||         ||||||::::::   |||||||||   |||
Db    2201 CGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 2260

Qy     162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178
              ||||||||   ::::::|||:::|||:::   ||||||      |||
Db    2261 GATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATGCAGACAGCAGGA 2320

Qy     179 ---------AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
                    :::::::::   |||   ::::::   :::|||      :::   |||
Db    2321 AATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATTGATGCCTTGGAA 2380

Qy     196 ArgLysThrAsnSerLeuProLeuHisSerVal------------------LysSerIle 209
              :::|||         :::                  :::|||:::
Db    2381 ATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGTC 2440

Qy     210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228
           :::   |||         |||:::      :::   ::::::   :::   |||:::
Db    2441 CGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTGGATGAAAGTATA 2500

Qy     229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
           |||   ||||||         :::|||:::||||||||||      ||||||||||:::   |||
Db    2501 CCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCCTTATTTTGTT 2560

Qy     249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
           :::|||         :::::::::||||||   :::   |||   |||:::|||||||||
Db    2561 GTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTTGCAATAATATTT 2620

Qy     269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
           :::|||||||||   :::|||         :::|||         |||   :::   ::::::
Db    2621 TCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAACGACAGAATAGT 2680

Qy     288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
           ::::::|||::::::|||:::   ||||||:::|||   |||:::::::::|||:::
Db    2681 AACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATTACATTTTTCCTT 2740

Qy     308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
           ||||||      :::|||:::||||||||||||||||||   |||||||||::::::
Db    2741 CAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTT 2800
```

```
Qy    328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
Db   2801 TTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAAAACACCACTGGA 2860

Qy    348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db   2861 GCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGG 2920

Qy    368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db   2921 CTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATC 2980

Qy    388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db   2981 TATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGA 3040

Qy    408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db   3041 GTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGT 3100

Qy    428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db   3101 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAG 3160

Qy    448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db   3161 GAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 3220

Qy    468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db   3221 AGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCC 3280

Qy    488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db   3281 TATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAG 3340

Qy    508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db   3341 GATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGT 3400

Qy    528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db   3401 TCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATT 3460

Qy    548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db   3461 GAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAA 3520

Qy    568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db   3521 GGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTG 3580

Qy    588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db   3581 CTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGT 3640

Qy    608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db   3641 GGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGG 3700

Qy    628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
Db   3701 AAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACAC 3760

Qy    648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db   3761 CTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCC 3820

Qy    668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db   3821 TATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCC 3880

Qy    688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db   3881 AACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAA 3940

Qy    708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
Db   3941 GGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGA 4000

Qy    728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db   4001 CAGCCTCATATTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAG 4060

Qy    748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db   4061 GTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCAC 4120

Qy    768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db   4121 CGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGCAGAGTCAAG 4180

Qy    788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db   4181 GAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGT 4240

Qy    808 AlaGlnSer 810
Db   4241 GTCCAGGCT 4249
```

RESULT 5
US-08-793-610-5
; Sequence 5, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
; APPLICANT: BAUM, Christopher
; APPLICANT: STOCKING-HARBERS, Carol
; APPLICANT: OSTERTAG, Wolfram
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEROF
; TITLE OF INVENTION: FOR GENE TRANSFER
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,610
; FILING DATE: 07-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 31 973.8
; FILING DATE: 08-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 03 952.1
; FILING DATE: 07-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/03175
; FILING DATE: 10-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berman, Richard J.
; REGISTRATION NUMBER: 39,105
; REFERENCE/DOCKET NUMBER: P1614-7007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6505 base pairs
; TYPE: nucleic acid

```
;     STRANDEDNESS:  double
;     TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA
US-08-793-610-5

Alignment Scores:
Pred. No.:                1.45e-274     Length:          6505
Score:                    2436.00       Matches:         466
Percent Similarity:       76.79%        Conservative:    166
Best Local Similarity:    56.62%        Mismatches:      177
Query Match:              59.72%        Indels:          14
DB:                       2             Gaps:            4

US-09-873-409-2 (1-812) x US-08-793-610-5 (1-6505)

Qy        2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db     3173 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 3232

Qy       22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db     3233 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 3292

Qy       42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db     3293 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 3352

Qy       62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db     3353 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 3412

Qy       82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db     3413 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 3472

Qy      102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db     3473 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 3532

Qy      122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db     3533 CTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGTTTGTCTACAGTT 3592

Qy      142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db     3593 CGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 3652

Qy      162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178
Db     3653 GATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATGCAGACAGCAGGA 3712

Qy      179 ---------AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
Db     3713 AATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATTGATGCCTTGGAA 3772

Qy      196 ArgLysThrAsnSerLeuProLeuHisSerVal------------------LysSerIle 209
Db     3773 ATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGTC 3832

Qy      210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228
Db     3833 CGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTGGATGAAAGTATA 3892

Qy      229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
Db     3893 CCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCCTTATTTTGTT 3952

Qy      249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db     3953 GTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTTGCAATAATATTT 4012

Qy      269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
Db     4013 TCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAACGACAGAATAGT 4072

Qy      288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db     4073 AACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATTACATTTTTCCTT 4132

Qy      308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db     4133 CAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTT 4192

Qy      328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
Db     4193 TTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAAAACACCACTGGA 4252

Qy      348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db     4253 GCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGG 4312

Qy      368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db     4313 CTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATC 4372

Qy      388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db     4373 TATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGA 4432

Qy      408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db     4433 GTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGT 4492

Qy      428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db     4493 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAG 4552

Qy      448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db     4553 GAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 4612

Qy      468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db     4613 AGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCC 4672

Qy      488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db     4673 TATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAG 4732

Qy      508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db     4733 GATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGT 4792

Qy      528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db     4793 TCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATT 4852

Qy      548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db     4853 GAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAA 4912

Qy      568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db     4913 GGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTG 4972

Qy      588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db     4973 CTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGT 5032

Qy      608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db     5033 GGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGG 5092

Qy      628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
Db     5093 AAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACAC 5152
```

```
Qy     648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db    5153 CTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCC 5212

Qy     668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db    5213 TATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCC 5272

Qy     688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db    5273 AACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAA 5332

Qy     708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
Db    5333 GGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGA 5392

Qy     728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db    5393 CAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAG 5452

Qy     748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db    5453 GTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCAC 5512

Qy     768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db    5513 CGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGCAGAGTCAAG 5572

Qy     788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db    5573 GAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGT 5632

Qy     808 AlaGlnSer 810
Db    5633 GTCCAGGCT 5641
```

RESULT 6
US-08-793-610-6
; Sequence 6, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
; APPLICANT: BAUM, Christopher
; APPLICANT: STOCKING-HARBERS, Carol
; APPLICANT: OSTERTAG, Wolfram
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEROF
; TITLE OF INVENTION: FOR GENE TRANSFER
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,610
; FILING DATE: 07-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 31 973.8
; FILING DATE: 08-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 03 952.1
; FILING DATE: 07-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/03175
; FILING DATE: 10-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berman, Richard J.
; REGISTRATION NUMBER: 39,105
; REFERENCE/DOCKET NUMBER: P1614-7007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9318 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA
US-08-793-610-6

Alignment Scores:

Pred. No.:	2.72e-274	Length:	9318
Score:	2436.00	Matches:	466
Percent Similarity:	76.79%	Conservative:	166
Best Local Similarity:	56.62%	Mismatches:	177
Query Match:	59.72%	Indels:	14
DB:	2	Gaps:	4

US-09-873-409-2 (1-812) x US-08-793-610-6 (1-9318)

```
Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db    3132 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 3191

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db    3192 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 3251

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db    3252 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 3311

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db    3312 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 3371

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db    3372 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 3431

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db    3432 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 3491

Qy     122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db    3492 CTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGTTTGTCTACAGTT 3551

Qy     142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db    3552 CGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 3611

Qy     162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178
Db    3612 GATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATGCAGACAGCAGGA 3671

Qy     179 ---------AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
Db    3672 AATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATTGATGCCTTGGAA 3731

Qy     196 ArgLysThrAsnSerLeuProLeuHisSerVal------------------LysSerIle 209
Db    3732 ATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGTC 3791

Qy     210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228
Db    3792 CGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTGGATGAAAGTATA 3851
```

```
Qy      229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
            |||   ||||||         :::|||:::||||||||||      |||||||||:::    |||
Db     3852 CCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCCTTATTTTGTT 3911

Qy      249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
            :::|||         ::::::::::||||||      :::      |||    |||:::|||||||||
Db     3912 GTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTTGCAATAATATTT 3971

Qy      269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
            :::|||||||||   :::|||            :::|||      |||   :::   ::::::
Db     3972 TCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAACGACAGAATAGT 4031

Qy      288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
               ::::::|||::::::|||:::      ||||||:::|||   |||:::::::::|||:::
Db     4032 AACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATTACATTTTTCCTT 4091

Qy      308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
            ||||||        :::|||:::||||||||||||||||||||||   |||||||||::::::
Db     4092 CAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTT 4151

Qy      328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
            |||::::::||||||   ||||||::::::||||||:::      :::|||:::||||||
Db     4152 TTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAAAACACCACTGGA 4211

Qy      348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
               |||||||||   ||||||   |||   ||||||::::::||||||   ||||||||
Db     4212 GCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGG 4271

Qy      368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
            :::      |||:::|||||||||      |||:::|||      :::|||||||||||||||
Db     4272 CTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATC 4331

Qy      388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
            ||||||||::::::|||   |||:::|||:::|||   |||::::::|||:::    |||
Db     4332 TATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGA 4391

Qy      408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
            ::::::|||         ::::::|||   |||      |||||||||:::||||||:::
Db     4392 GTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGT 4451

Qy      428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
            ||||||||||||||||||||||||:::||||||   ||||||:::||||||||||||:::
Db     4452 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAG 4511

Qy      448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
            |||:::   ||||||   ||||||   :::   ||||||      :::||||||:::
Db     4512 GAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 4571

Qy      468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
            :::||||||   |||   |||      ::::::|||:::   |||   :::||||||:::
Db     4572 AGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCC 4631

Qy      488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
            ||||||   |||||||||||||||||||||:::            |||:::   |||
Db     4632 TATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAG 4691

Qy      508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
               :::   :::||||||:::|||:::   :::||||||||||||:::|||:::
Db     4692 GATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGT 4751

Qy      528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
                     |||||:::|||:::|||||||||      |||||||||:::      ::::::
Db     4752 TCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATT 4811

Qy      548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
            |||||   |||   |||||||||   |||   ||||||      |||:::|||   |||
Db     4812 GAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAA 4871

Qy      568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
            ||||||:::   |||   ||||||   |||   ||||||   |||||||||:::    :::
Db     4872 GGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTG 4931

Qy      588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
            |||:::||||||||||||      ::::::::|||:::|||:::|||   ||||||||||||
Db     4932 CTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGT 4991

Qy      608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
            ||||||||||||||||||   |||||||||||:::|||   |||||||||:::    |||
Db     4992 GGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGG 5051

Qy      628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
            :::||||||   ||||||   :::   |||   ||||||||||||||||||||||:::
Db     5052 AAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACAC 5111

Qy      648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
            :::   ||||||   |||||||||:::||||||:::|||||||||||||||||||||||||
Db     5112 CTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCC 5171

Qy      668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
            ||||||||||||||||||||||         :::||||||      ||||||      |||
Db     5172 TATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCC 5231

Qy      688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
            |||||||||:::|||||||||   ||||||   ||||||:::|||:::||||||   |||
Db     5232 AACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAA 5291

Qy      708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
            |||   |||||||||||||||||||||||||:::|||||||||||||||||||||::::::
Db     5292 GGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGA 5351

Qy      728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
            :::|||   |||||||||||||||||||||||||||||||||      :::|||||||||
Db     5352 CAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAG 5411

Qy      748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
            |||||||||   ||||||||||||||||||   ||||||||||||:::|||:::    |||
Db     5412 GTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCAC 5471

Qy      768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
            |||||||||   ||||||||||||||||||||||||||||         ||||||::::::|||
Db     5472 CGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGCAGAGTCAAG 5531

Qy      788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
            |||   ||||||||||||:::||||||      :::   |||||||||      :::|||:::
Db     5532 GAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGT 5591

Qy      808 AlaGlnSer 810
               |||:::
Db     5592 GTCCAGGCT 5600
```

RESULT 7
US-09-120-513-1
; Sequence 1, Application US/09120513
; Patent No. 6025160
; GENERAL INFORMATION:
;   APPLICANT: Brun, Kimberly
;   APPLICANT: Chenery, Richard
;   APPLICANT: Ellens, Harma
;   APPLICANT: Field, John
;   APPLICANT: Yue, Lin
;   TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE
;   TITLE OF INVENTION: SEQUENCES ENCODING RAT MDR1B2 AND
;   TITLE OF INVENTION: SCREENING METHODS THEREOF
;   NUMBER OF SEQUENCES: 2
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: SmithKline Beecham Corporation
;   STREET: 709 Swedeland Road
;   CITY: King of Prussia
;   STATE: PA
;   COUNTRY:
;   ZIP: 19406
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ for Windows Version 2.0

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/120,513
;      FILING DATE:  22-JUL-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  King, William T
;      REGISTRATION NUMBER:  30,954
;      REFERENCE/DOCKET NUMBER:  GP50008
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-5015
;      TELEFAX:  610-270-5090
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4233 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
US-09-120-513-1

Alignment Scores:
Pred. No.:              2.97e-272     Length:        4233
Score:                  2413.50       Matches:       467
Percent Similarity:     76.03%        Conservative:  158
Best Local Similarity:  56.81%        Mismatches:    180
Query Match:            59.17%        Indels:        17
DB:                     3             Gaps:          5

US-09-873-409-2 (1-812) x US-09-120-513-1 (1-4233)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db   1376 ATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTG 1435

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db   1436 GTGAGTCAGGAACCCGTGCTGTTTGCCACCACGATTGCCGAAAACATTCGCTATGGCCGA 1495

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db   1496 GAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTC 1555

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db   1556 ATCATGAAACTGCCCCACAAATTTAACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGT 1615

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db   1616 GGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAACCCCAAGATCCTT 1675

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db   1676 TTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTGGTTCAGGCCGCT 1735

Qy    122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db   1736 CTGGATAAGGCTAGAGAAGGCCGGACCACCATTGTGATAGCTCACCGCTTGTCTACAGTG 1795

Qy    142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db   1796 CGCAATGCTGACGTCATTGCTGGTTTTGATGGTGGTGTCATTGTGGAGCAAGGAAATCAT 1855

Qy    162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLys 181
Db   1856 GAAGAGCTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTCATGACACAG---ACTAGA 1912

Qy    182 LysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu 201
Db   1913 GGAAATGAAATTGAACCAGGAAATAATGCTTATGAATCCCAAAGTGACACTGGTGCCTCT 1972

Qy    202 ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAla------------ 217
Db   1973 GAGTTGACTTCAGAAAAATCA---AAATCTCCTTTAATAAGGAGATCAATTCGCAGAAGT 2029

Qy    218 --------------------GluGluSerThrGlnSerLysGlu---------IleSer 227
Db   2030 ATCCACAGAAGACAAGACCAGGAGAGAAGACTTAGTTCGAAAGAGGATGTGGATGAAGAT 2089

Qy    228 LeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheVal 247
Db   2090 GTGCCTATGGTTTCCTTTTGGCAGATCCTAAAGCTAAATATTAGTGAATGGCCCTATTTA 2149

Qy    248 ValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIle 267
Db   2150 GTTGTGGGTGTACTTTGTGCTGTTATAAATGGGTGCATACAACCAGTGTTTGCCATAGTG 2209

Qy    268 PheAlaLysIleIleThrMetPheGly---AsnAsnAspLysThrThrLeuLysHisAsp 286
Db   2210 TTTTCAAAGATTGTAGGGGTTTTTTCAAGAGACGACGACCATGAAACCAAACAACGGAAT 2269

Qy    287 AlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPhe 306
Db   2270 TGTAACTTGTTTTCCCTTCTCTTTCTGGTCATGGGAATGATTTCTTTTGTTACGTACTTC 2329

Qy    307 MetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeu 326
Db   2330 TTTCAAGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGACTCCGATACATG 2389

Qy    327 AlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThr 346
Db   2390 GTCTTCAAATCCATGCTGCGACAGGATATAAGCTGGTTTGATGACCATAAAAACACCACT 2449

Qy    347 GlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySer 366
Db   2450 GGCTCGCTGACTACCAGGCTCGCTAGTGACGCTTCTAATGTTAAAGGGGCTATGGGCTCC 2509

Qy    367 ArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPhe 386
Db   2510 AGGCTTGCTGTAGTTACCCAGAATGTAGCAAACCTTGGCACAGGAATTATCTTATCCTTA 2569

Qy    387 IleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThr 406
Db   2570 GTCTATGGCTGGCAGCTTACACTTTTACTTGTAGTAATTATACCACTCATTGTCTTGGGT 2629

Qy    407 GlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLys 426
Db   2630 GGAATTATTGAAATGAAACTGTTGTCTGGTCAAGCCTTGAAGGACAAGAAAGAGCTAGAG 2689

Qy    427 HisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThr 446
Db   2690 ATCTCTGGGAAGATCGCTACAGAAGCAATTGAAAACTTCCGCACTGTTGTCTCTTTGACT 2749

Qy    447 ArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThr 466
Db   2750 AGGGAGCAGAAGTTTGAAACTATGTATGCCCAGAGCTTGCAGATACCATACAGAAATGCT 2809

Qy    467 SerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPhe 486
Db   2810 TTGAAGAAAGCACACGTCTTTGGGATCACCTTCGCCTTCACCCAGGCCATGATTTATTTT 2869

Qy    487 AlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrPro 506
Db   2870 TCCTATGCTGCTTGTTTCCGGTTCGGTGCCTACTTGGTGGCACGAGAACTCATGACGTTT 2929

Qy    507 GluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThr 526
Db   2930 GAAAATGTTATGTTGGTATTTTCTGCTGTTGTCTTTGGTGCCATGGCAGCAGGGAATACC 2989

Qy    527 LeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeu 546
Db   2990 AGTTCATTCGCTCCTGACTACGCGAAGGCCAAAGTCTCAGCATCCCACATCATCAGGATC 3049

Qy    547 LeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCys 566
Db   3050 ATTGAGAAAATCCCCGAGATTGACAGCTACAGCACGGAGGGCTTGAAGCCTAATTGGTTA 3109
```

```
Qy     567 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 586
           |||||||::::::|||      |||   |||   |||||   |||||::::::
Db    3110 GAAGGAAATGTGAAATTTAATGGAGTCATGTTCAACTATCCCACCCGACCCAACATCCCA 3169

Qy     587 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 606
           :::|||:::|||||||     :::::::::|||:::|||:::      ||||||||
Db    3170 GTGCTTCAGGGACTGAGCTTCGAGGTGAAGAAGGGGCAAACGCTTCGCCTGGTGGGCAGC 3229

Qy     607 SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGln 626
           ||||||||||||||||||   |||||||||||:::|||    |||:::|||:::
Db    3230 AGTGGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTCGAGCGCTTCTACAACCCCATGGCT 3289

Qy     627 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 646
           |||   |||       |||||   :::   |||:::||||||||||||||||||:::
Db    3290 GGAACAGTGTTTCTAGATGGCAAAGAAATAAAACAACTCAACGTCCAGTGGCTCCGCGCC 3349

Qy     647 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 666
              :::    ||||||   |||||||||:::||||||:::||||||||   ||||||||
Db    3350 CACCTGGGCATTGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATCACCGAGAACATC 3409

Qy     667 AlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla 686
           ||||||||||||||||||||||||       :::||||||     ||||||
Db    3410 GCCTACGGAGACAACAGCCGTGTCGTGTCTCATGAGGAGATCGTGAGGGCCGCCAGGGAG 3469

Qy     687 AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeu 706
           ||||||||||||   ||||||:::   |||||||||||||||||||||:::||||||
Db    3470 GCCAACATCCACCAGTTCATCGACTCACTGCCTGAGAAATACAACACCAGAGTGGGAGAC 3529

Qy     707 LysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeu 726
           ||||||   ||||||||||||||||||||||||:::||||||||||||||||||:::
Db    3530 AAAGGGACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATCGCCATCGCGCGCGCCCTCGTC 3589

Qy     727 GlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGlu 746
           ::::::|||   |||||||||||||||||||||||||||||       :::||||||
Db    3590 AGACAGCCTCACATCTTACTTCTGGATGAAGCGACATCAGCTCTGGATACGGAGAGTGAA 3649

Qy     747 LysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThr 766
           ||||||||||||   ||||||||||||||||||   ||||||||||:::|||:::
Db    3650 AAGGTCGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACCTGCATTGTGATCGCG 3709

Qy     767 HisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIle 786
           ||||||||||||   |||||||||||||||||||||||||||:::   ||||||::::::
Db    3710 CACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATTCAGAACGGCCAGGTC 3769

Qy     787 LysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuVal 806
           ||||||   ||||||||||||:::||||||        :::   |||||||||   :::|||
Db    3770 AAGGAGCACGGCACCCACCAGCAGCTGCTGGCCCAGAAAGGCATCTATTTCTCGATGGTT 3829

Qy     807 AsnAla 808
              |||
Db    3830 CAGGCT 3835
```

RESULT 8
US-09-450-105-1
; Sequence 1, Application US/09450105
; Patent No. 6169166
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: ENCODING RAT MDR1B2 AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP-50008-D1
; CURRENT APPLICATION NUMBER: US/09/450,105
; CURRENT FILING DATE: 1999-11-29
; EARLIER APPLICATION NUMBER: 09/120,513
; EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 4233
;   TYPE: DNA
;   ORGANISM: HOMO SAPIENS
US-09-450-105-1

Alignment Scores:

Pred. No.:	2.97e-272	Length:	4233
Score:	2413.50	Matches:	467
Percent Similarity:	76.03%	Conservative:	158
Best Local Similarity:	56.81%	Mismatches:	180
Query Match:	59.17%	Indels:	17
DB:	4	Gaps:	5

US-09-873-409-2 (1-812) x US-09-450-105-1 (1-4233)

```
Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
           :::|||      ||||||||   :::||||||||:::   |||:::   ||||||||
Db    1376 ATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTG 1435

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           ||||||||||||||||||||||||   |||||||||:::   ||||||:::||||||||
Db    1436 GTGAGTCAGGAACCCGTGCTGTTTGCCACCACGATTGCCGAAAACATTCGCTATGGCCGA 1495

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
           ::::::||||||   :::|||:::|||:::|||   :::|||||||||||||||||||||
Db    1496 GAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTC 1555

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           ||||||:::   |||:::||||||||||||||||||||||||:::|||||||||:::|||
Db    1556 ATCATGAAACTGCCCCACAAATTTAACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGT 1615

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1616 GGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAACCCCAAGATCCTT 1675

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
           :::||||||||||||||||||||||||||||:::||||||::::::   ||||||||||||
Db    1676 TTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTGGTTCAGGCCGCT 1735

Qy     122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
           |||:::||||||   :::||||||||||||||||||||||||||||||||||||||:::
Db    1736 CTGGATAAGGCTAGAGAAGGCCGGACCACCATTGTGATAGCTCACCGCTTGTCTACAGTG 1795

Qy     142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
           |||:::||||||:::|||               |||::::::   |||:::|||   |||
Db    1796 CGCAATGCTGACGTCATTGCTGGTTTTGATGGTGGTGTCATTGTGGAGCAAGGAAATCAT 1855

Qy     162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLys 181
              |||||||||   ::::::|||:::|||:::   |||||||||:::|||      :::
Db    1856 GAAGAGCTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTCATGACACAG---ACTAGA 1912

Qy     182 LysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu 201
                 :::   :::      :::      |||   ::::::      |||   :::
Db    1913 GGAAATGAAATTGAACCAGGAAATAATGCTTATGAATCCCAAAGTGACACTGGTGCCTCT 1972

Qy     202 ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAla------------ 217
           |||   |||   ||||||   ||||||      |||   ::::::
Db    1973 GAGTTGACTTCAGAAAAATCA---AAATCTCCTTTAATAAGGAGATCAATTCGCAGAAGT 2029

Qy     218 --------------------GluGluSerThrGlnSerLysGlu---------IleSer 227
                               |||         ||||||||
Db    2030 ATCCACAGAAGACAAGACCAGGAGAGAAGACTTAGTTCGAAAGAGGATGTGGATGAAGAT 2089

Qy     228 LeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheVal 247
           :::|||   ||||||      :::||||||||||||||      ||||||||::::::
Db    2090 GTGCCTATGGTTTCCTTTTGGCAGATCCTAAAGCTAAATATTAGTGAATGGCCCTATTTA 2149

Qy     248 ValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIle 267
           |||:::|||   |||   :::|||:::||||||   :::   |||||||||:::|||:::
Db    2150 GTTGTGGGTGTACTTTGTGCTGTTATAAATGGGTGCATACAACCAGTGTTTGCCATAGTG 2209

Qy     268 PheAlaLysIleIleThrMetPheGly---AsnAsnAspLysThrThrLeuLysHisAsp 286
           |||:::||||||:::   :::|||      ::::::|||      |||   :::   :::
```

```
Db   2210 TTTTCAAAGATTGTAGGGGTTTTTCAAGAGACGACGACCATGAAACCAAACAACGGAAT 2269

Qy    287 AlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPhe 306
Db   2270 TGTAACTTGTTTTCCCTTCTCTTTCTGGTCATGGGAATGATTTCTTTTGTTACGTACTTC 2329

Qy    307 MetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeu 326
Db   2330 TTTCAAGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGACTCCGATACATG 2389

Qy    327 AlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThr 346
Db   2390 GTCTTCAAATCCATGCTGCGACAGGATATAAGCTGGTTTGATGACCATAAAAACACCACT 2449

Qy    347 GlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySer 366
Db   2450 GGCTCGCTGACTACCAGGCTCGCTAGTGACGCTTCTAATGTTAAAGGGGCTATGGGCTCC 2509

Qy    367 ArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPhe 386
Db   2510 AGGCTTGCTGTAGTTACCCAGAATGTAGCAAACCTTGGCACAGGAATTATCTTATCCTTA 2569

Qy    387 IleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThr 406
Db   2570 GTCTATGGCTGGCAGCTTACACTTTTACTTGTAGTAATTATACCACTCATTGTCTTGGGT 2629

Qy    407 GlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLys 426
Db   2630 GGAATTATTGAAATGAAACTGTTGTCTGGTCAAGCCTTGAAGGACAAGAAAGAGCTAGAG 2689

Qy    427 HisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThr 446
Db   2690 ATCTCTGGGAAGATCGCTACAGAAGCAATTGAAAACTTCCGCACTGTTGTCTCTTTGACT 2749

Qy    447 ArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThr 466
Db   2750 AGGGAGCAGAAGTTTGAAACTATGTATGCCCAGAGCTTGCAGATACCATACAGAAATGCT 2809

Qy    467 SerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPhe 486
Db   2810 TTGAAGAAAGCACACGTCTTTGGGATCACCTTCGCCTTCACCCAGGCCATGATTTATTTT 2869

Qy    487 AlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrPro 506
Db   2870 TCCTATGCTGCTTGTTTCCGGTTCGGTGCCTACTTGGTGGCACGAGAACTCATGACGTTT 2929

Qy    507 GluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThr 526
Db   2930 GAAAATGTTATGTTGGTATTTTCTGCTGTTGTCTTTGGTGCCATGGCAGCAGGGAATACC 2989

Qy    527 LeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeu 546
Db   2990 AGTTCATTCGCTCCTGACTACGCGAAGGCCAAAGTCTCAGCATCCCACATCATCAGGATC 3049

Qy    547 LeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCys 566
Db   3050 ATTGAGAAAATCCCCGAGATTGACAGCTACAGCACGGAGGGCTTGAAGCCTAATTGGTTA 3109

Qy    567 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 586
Db   3110 GAAGGAAATGTGAAATTTAATGGAGTCATGTTCAACTATCCCACCCGACCCAACATCCCA 3169

Qy    587 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 606
Db   3170 GTGCTTCAGGGACTGAGCTTCGAGGTGAAGAAGGGGCAAACGCTTCGCCTGGTGGGCAGC 3229

Qy    607 SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGln 626
Db   3230 AGTGGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTCGAGCGCTTCTACAACCCCATGGCT 3289

Qy    627 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 646
Db   3290 GGAACAGTGTTTCTAGATGGCAAAGAAATAAAACAACTCAACGTCCAGTGGCTCCGCGCC 3349

Qy    647 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 666
Db   3350 CACCTGGGCATTGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATCACCGAGAACATC 3409

Qy    667 AlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla 686
Db   3410 GCCTACGGAGACAACAGCCGTGTCGTGTCTCATGAGGAGATCGTGAGGGCCGCCAGGGAG 3469

Qy    687 AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeu 706
Db   3470 GCCAACATCCACCAGTTCATCGACTCACTGCCTGAGAAATACAACACCAGAGTGGGAGAC 3529

Qy    707 LysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeu 726
Db   3530 AAAGGGACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATCGCCATCGCGCGCGCCCTCGTC 3589

Qy    727 GlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGlu 746
Db   3590 AGACAGCCTCACATCTTACTTCTGGATGAAGCGACATCAGCTCTGGATACGGAGAGTGAA 3649

Qy    747 LysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThr 766
Db   3650 AAGGTCGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACCTGCATTGTGATCGCG 3709

Qy    767 HisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIle 786
Db   3710 CACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATTCAGAACGGCCAGGTC 3769

Qy    787 LysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuVal 806
Db   3770 AAGGAGCACGGCACCCACCAGCAGCTGCTGGCCCAGAAAGGCATCTATTTCTCGATGGTT 3829

Qy    807 AsnAla 808
Db   3830 CAGGCT 3835
```

RESULT 9
US-08-752-447-1
; Sequence 1, Application US/08752447
; Patent No. 5994088
; GENERAL INFORMATION:
; APPLICANT: Mechetner, Eugene
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods and Reagents for Preparing and
; TITLE OF INVENTION: Using Immunoligcal Agents Specific for P-glycoprotein
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff Ltd.
; STREET: 300 South Wacker Drive, Seventh Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/752,447
; FILING DATE: 15-NOV-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5994088nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 95,1121
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-9808
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4669 base pairs

```
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  5'UTR
;      LOCATION:  1..424
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  425..4264
;    FEATURE:
;      NAME/KEY:  3'UTR
;      LOCATION:  4265..4669
US-08-752-447-1

Alignment Scores:
Pred. No.:               5.29e-272     Length:         4669
Score:                   2412.00       Matches:        462
Percent Similarity:      76.06%        Conservative:   164
Best Local Similarity:   56.14%        Mismatches:     183
Query Match:             59.13%        Indels:         14
DB:                      2             Gaps:           4

US-09-873-409-2 (1-812) x US-08-752-447-1 (1-4669)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db    1781 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 1840

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db    1841 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 1900

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db    1901 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1960

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db    1961 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 2020

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db    2021 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 2080

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db    2081 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 2140

Qy     122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db    2141 CTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGTTTTGCTACAGTT 2200

Qy     142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db    2201 CGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 2260

Qy     162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178
Db    2261 GATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATGCAGACAGCAGGA 2320

Qy     179 ---------AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
Db    2321 AATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATTGATGCCTTGGAA 2380

Qy     196 ArgLysThrAsnSerLeuProLeuHisSerVal------------------LysSerIle 209
Db    2381 ATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGTC 2440

Qy     210 LysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLysGluIleSerLeu 228
Db    2441 CGTGGATCACAAGCCCAACACAGAAAGCTTAGTACCAAAGAGGCTCTGGATGAAAGTATA 2500

Qy     229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
Db    2501 CCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCCTTATTTTGTT 2560

Qy     249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db    2561 GTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTTGCAATAATATTT 2620

Qy     269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
Db    2621 TCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAACGACAGAATAGT 2680

Qy     288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db    2681 AACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATTACATTTTTCCTT 2740

Qy     308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db    2741 CAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTT 2800

Qy     328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
Db    2801 TTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTCATGACCCTAAAAACACCACTGGA 2860

Qy     348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db    2861 GCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGG 2920

Qy     368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387 }
Db    2921 CTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATC 2980

Qy     388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db    2981 TATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGA 3040

Qy     408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db    3041 GTTGTTGAAATGAAAATGTTTGCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGT 3100

Qy     428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db    3101 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAG 3160

Qy     448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db    3161 GAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 3220

Qy     468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db    3221 AGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCC 3280

Qy     488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db    3281 TATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAG 3340

Qy     508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db    3341 GATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGT 3400

Qy     528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db    3401 TCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATT 3460

Qy     548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db    3461 GAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAA 3520

Qy     568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db    3521 GGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTG 3580

Qy     588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
```

```
                |||:::|||||||||||  :::::::::|||:::|||:::|||  |||||||||||
Db       3581 CTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGT 3640

Qy        608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
              |||||||||||||||||  ||||||||||||:::|||  |||||||||:::   |||
Db       3641 GGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGG 3700

Qy        628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
              :::||||||  ||||||   :::   |||  ||||||||||||||||||||||:::
Db       3701 AAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACAC 3760

Qy        648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
              :::   ||||||  ||||||||:::||||||:::||||||||||||||||||||||||
Db       3761 CTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCC 3820

Qy        668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
              |||||||||||||||||||||||      :::||||||      ||||||      |||
Db       3821 TATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCC 3880

Qy        688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
              |||||||||:::|||||||||   ||||||  ||||||:::|||:::||||||   |||
Db       3881 AACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAA 3940

Qy        708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
              |||   |||||||||||||||||||||||||:::||||||||||||   |||::::::
Db       3941 GGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTCGCCTTGTTAGA 4000

Qy        728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
              :::|||   ||||||||||||||||||||||||||||||||||||   :::|||||||||
Db       4001 CAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAG 4060

Qy        748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
              |||||||||   ||||||||||||||||||   ||||||||||||:::|||:::   |||
Db       4061 GTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCAC 4120

Qy        768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
              |||||||||   |||||||||||||||||||||||||||         ||||||::::::|||
Db       4121 CGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGCAGAGTCAAG 4180

Qy        788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
              |||   |||||||||||||:::||||||      :::   |||||||||   :::|||:::
Db       4181 GAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGT 4240

Qy        808 AlaGlnSer 810
                 |||:::
Db       4241 GTCCAGGCT 4249
```

RESULT 10
US-09-316-167-1
; Sequence 1, Application US/09316167
; Patent No. 6365357
; GENERAL INFORMATION:
; APPLICANT: Mechetner, Eugene
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods and Reagents for Preparing and
; TITLE OF INVENTION: Using Immunoligcal Agents Specific for P-glycoprotein
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff Ltd.
; STREET: 300 South Wacker Drive, Seventh Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/316,167
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/752,447
; FILING DATE: 15-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6365357nan; Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 95,1121
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-9808
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..424
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 425..4264
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 4265..4669
US-09-316-167-1

```
Alignment Scores:
Pred. No.:               5.29e-272     Length:        4669
Score:                   2412.00       Matches:       462
Percent Similarity:      76.06%        Conservative:  164
Best Local Similarity:   56.14%        Mismatches:    183
Query Match:             59.13%        Indels:        14
DB:                      4             Gaps:          4
```

US-09-873-409-2 (1-812) x US-09-316-167-1 (1-4669)

```
Qy          2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
              ||||||     |||||||||   :::|||||||||||   ||||:::   |||||||||
Db       1781 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 1840

Qy         22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
              |||||||||||||||||||||||||||   ||||||||:::   ||||||:::|||||||||
Db       1841 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 1900

Qy         42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
              ::::::||||||   :::|||:::|||:::|||   :::|||||||||||||||||||||
Db       1901 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1960

Qy         62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
              ||||||:::   |||:::||||||:::|||||||||||||||:::|||||||||:::|||
Db       1961 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 2020

Qy         82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2021 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 2080

Qy        102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
              :::||||||||||||||||||||||||||||:::||||||::::::   ||||||   |||
Db       2081 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 2140

Qy        122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
              |||:::||||||   |||||||||||||||||||||:::|||||||||   :::|||:::
Db       2141 CTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGTTTGCTACAGTT 2200

Qy        142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
              |||:::||||||:::|||         ||||||::::::   |||||||||   |||
Db       2201 CGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 2260

Qy        162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178
                 |||||||||   ::::::|||:::|||:::   ||||||      |||
Db       2261 GATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATGCAGACAGCAGGA 2320
```

```
Qy    179 ---------AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
Db   2321 AATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATTGATGCCTTGGAA 2380

Qy    196 ArgLysThrAsnSerLeuProLeuHisSerVal------------------LysSerIle 209
Db   2381 ATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGTC 2440

Qy    210 LysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLysGluIleSerLeu 228
Db   2441 CGTGGATCACAAGCCCAACACAGAAAGCTTAGTACCAAAGAGGCTCTGGATGAAAGTATA 2500

Qy    229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
Db   2501 CCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCCTTATTTTGTT 2560

Qy    249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db   2561 GTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTTGCAATAATATTT 2620

Qy    269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
Db   2621 TCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAACGACAGAATAGT 2680

Qy    288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db   2681 AACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATTACATTTTTCCTT 2740

Qy    308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db   2741 CAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTT 2800

Qy    328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
Db   2801 TTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTCATGACCCTAAAAACACCACTGGA 2860

Qy    348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db   2861 GCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGG 2920

Qy    368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db   2921 CTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATC 2980

Qy    388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db   2981 TATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGA 3040

Qy    408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db   3041 GTTGTTGAAATGAAAATGTTTGCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGT 3100

Qy    428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db   3101 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAG 3160

Qy    448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db   3161 GAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 3220

Qy    468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db   3221 AGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCC 3280

Qy    488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db   3281 TATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAG 3340

Qy    508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db   3341 GATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGT 3400

Qy    528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db   3401 TCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATT 3460

Qy    548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db   3461 GAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAA 3520

Qy    568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db   3521 GGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTG 3580

Qy    588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db   3581 CTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGT 3640

Qy    608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db   3641 GGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGG 3700

Qy    628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
Db   3701 AAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACAC 3760

Qy    648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db   3761 CTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCC 3820

Qy    668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db   3821 TATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCC 3880

Qy    688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db   3881 AACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAA 3940

Qy    708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
Db   3941 GGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGA 4000

Qy    728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db   4001 CAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAG 4060

Qy    748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db   4061 GTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCAC 4120

Qy    768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db   4121 CGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGCAGAGTCAAG 4180

Qy    788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db   4181 GAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGT 4240

Qy    808 AlaGlnSer 810
Db   4241 GTCCAGGCT 4249
```

RESULT 11
US-08-583-276-18
; Sequence 18, Application US/08583276
; Patent No. 5837536
; GENERAL INFORMATION:
; APPLICANT: McDonagh, Kevin T.
; APPLICANT: Nienhuis, Arthur
; APPLICANT: Tolstoshev, Paul
; TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN
; TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED
; TITLE OF INVENTION: SELECTION OF CELLS TRANSDUCED WITH SUCH GENES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:

```
;       ADDRESSEE:  Carella, Byrne, Bain, Gilfillan,
;       ADDRESSEE:  Cecchi & Stewart
;       STREET:  6 Becker Farm Road
;       CITY:  Roseland
;       STATE:  New Jersey
;       COUNTRY:  USA
;       ZIP:  07068
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  3.5 inch diskette
;       COMPUTER:  IBM PS/2
;       OPERATING SYSTEM:  PC-DOS
;       SOFTWARE:  DW4.V2
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/583,276
;       FILING DATE:  05-JAN-1996
;       CLASSIFICATION:  435
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  08/332,444
;       FILING DATE:  31-OCT-1994
;       APPLICATION NUMBER:  07/887,712
;       FILING DATE:  22-MAY-1992
;   INFORMATION FOR SEQ ID NO:  18:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  4669 bases
;       TYPE:  nucleic acid
;       STRANDEDNESS:  singular
;       TOPOLOGY:  linear
;     MOLECULE TYPE:
;       DESCRIPTION:  Genomic DNA
US-08-583-276-18

Alignment Scores:
Pred. No.:              3.04e-270     Length:         4669
Score:                  2397.00       Matches:        460
Percent Similarity:     76.03%        Conservative:   165
Best Local Similarity:  55.96%        Mismatches:     183
Query Match:            58.76%        Indels:         14
DB:                     2             Gaps:           4

US-09-873-409-2 (1-812) x US-08-583-276-18 (1-4669)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db    1781 GTTCATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 1840

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db    1841 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 1900

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db    1901 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1960

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db    1961 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGCCCCAGTTGAGT 2020

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db    2021 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 2080

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db    2081 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 2140

Qy     122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db    2141 CTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGTTTGTCTACAGTT 2200

Qy     142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db    2201 CGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 2260

Qy     162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178
Db    2261 GATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATGCAGACAGCAGGA 2320

Qy     179 ---------AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
Db    2321 AATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATTGATGCCTTGGAA 2380

Qy     196 ArgLysThrAsnSerLeuProLeuHisSerVal------------------LysSerIle 209
Db    2381 ATGTCTTCAAATGATTCAAGATCACGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGTC 2440

Qy     210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228
Db    2441 CGTGGATCACAACCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTGGATGAAAGTATA 2500

Qy     229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
Db    2501 CCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCCTTATTTTGTT 2560

Qy     249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db    2561 GTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTTGCAATAATATTT 2620

Qy     269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
Db    2621 TCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAACGACAGAATACT 2680

Qy     288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db    2681 AACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATTACATTTTTCCTT 2740

Qy     308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db    2741 CAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTT 2800

Qy     328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
Db    2801 TTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAAAACACCACTGGA 2860

Qy     348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db    2861 GCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGG 2920

Qy     368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db    2921 CTTGCTGTAATTACCCAGAATATAGCAATTCTTGGGACAGGAATAATTATATCCTTCATC 2980

Qy     388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db    2981 TATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGA 3040

Qy     408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db    3041 GTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGT 3100

Qy     428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db    3101 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAG 3160

Qy     448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db    3161 GAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 3220

Qy     468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db    3221 AGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCC 3280

Qy     488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db    3281 TATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTAATGAGCTTTGAG 3340

Qy     508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
```

```
              :::   :::||||||:::|||:::   :::||||||||||:::|||:::
Db     3341 GATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTGAGT 3400

Qy      528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
                     ||||||:::|||:::||||||||      |||||||||:::      ::::::
Db     3401 TCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATT 3460

Qy      548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
            ||||||   |||   |||||||||   |||   ||||||         |||:::|||   |||
Db     3461 GAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAA 3520

Qy      568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
            ||||||:::   |||   ||||||   |||   ||||||   |||||||||:::   :::
Db     3521 GGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTC 3580

Qy      588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
            |||:::||||||||||||   :::::::::|||:::|||:::|||   ||||||||||||
Db     3581 CTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGT 3640

Qy      608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
            ||||||||||||||||||   ||||||||||||:::|||   |||||||||:::   |||
Db     3641 GGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGG 3700

Qy      628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
            :::||||||   ||||||   :::   |||   ||||||||||||||||||||||:::
Db     3701 AAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACAC 3760

Qy      648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
            :::   ||||||   |||||||||:::||||||:::||||||||||||||||||||||||
Db     3761 CTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCC 3820

Qy      668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
            ||||||   |||||||||||||||         :::||||||      ||||||     |||
Db     3821 TATGGACACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCC 3880

Qy      688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
            |||||||||:::|||||||||   ||||||   ||||||:::|||:::||||||   |||
Db     3881 AACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAA 3940

Qy      708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
            |||   |||||||||||||||||||||||||||:::||||||||||||||||||::::::
Db     3941 GGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGA 4000

Qy      728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
            :::|||   ||||||||||||||||||||||||||||||||||||   :::|||||||||
Db     4001 CAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAG 4060

Qy      748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
            ||||||||   ||||||||||||   |||   ||||||||||||:::|||:::   |||
Db     4061 GTTGTCCAAGAAGCCCTGGACAAACCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCAC 4120

Qy      768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
            ||||||||   ||||||||||||||||||||||||         ||||||::::::|||
Db     4121 CGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGCAGAGTCAAG 4180

Qy      788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
            :::   ||||||||||||:::||||||      :::   |||||||||   :::|||:::
Db     4181 CAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGT 4240

Qy      808 AlaGln 809
               |||
Db     4241 GTCCAG 4246
```

RESULT 12
US-08-461-823-1
; Sequence 1, Application US/08461823
; Patent No. 5593840
; GENERAL INFORMATION:
;   APPLICANT: Bhatnagar, Satish K.
;   APPLICANT: George Jr., Albert L.
;   APPLICANT: Nazarenko, Irina
;   TITLE OF INVENTION: AMPLIFICATION OF NUCLEIC ACID SEQUENCES
;   NUMBER OF SEQUENCES: 27
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OncorPharm, Inc.
;   STREET: 200 Perry Parkway
;   CITY: Gaithersburg
;   STATE: Maryland
;   COUNTRY: USA
;   ZIP: 20877
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/461,823
;   FILING DATE: 05-JUN-1995
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/168,621
;   FILING DATE: 16-DEC-1993
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/010,433
;   FILING DATE: 27-JAN-1993
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Karta, Glenn E.
;   REGISTRATION NUMBER: 30,649
;   REFERENCE/DOCKET NUMBER: PA-0012 CIP 2
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 301 527-2058
;   TELEFAX: 301 208-6997
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2726 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
US-08-461-823-1

```
Alignment Scores:
Pred. No.:              6.35e-257     Length:          2726
Score:                  2280.00       Matches:         435
Percent Similarity:     76.55%        Conservative:    159
Best Local Similarity:  56.06%        Mismatches:      168
Query Match:            55.90%        Indels:          14
DB:                     1             Gaps:            4

US-09-873-409-2 (1-812) x US-08-461-823-1 (1-2726)

Qy       49 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 68
            :::|||:::|||   :::||||||||||||||||||||||||||||:::   |||:::|||
Db        2 ATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAA 61

Qy       69 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 88
            |||:::|||||||||||||||:::|||||||||:::||||||||||||||||||||||||
Db       62 TTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATC 121

Qy       89 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 108
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      122 GCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCA 181

Qy      109 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 128
            ||||||||:::||||||:::::::   ||||||   ||||||:::||||||   ||||||
Db      182 GCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGT 241

Qy      129 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 148
            ||||||||||||:::||||||||||||||||:::|||:::||||||:::|||
Db      242 CGGACCACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCT 301

Qy      149 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 168
                     ||||||::::::   |||||||||   |||   |||||||||   ::::::
Db      302 GGTTTCGATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAA 361
```

```
Qy     169 GlyLeuTyrTyrSerLeuValMetSerGln------------------AspIleLysLys 182
Db     362 GGCATTTACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAAT 421

Qy     183 AlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPro 202
Db     422 GCAGCTGATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGA 481

Qy     203 LeuHisSerVal------------------LysSerIleLys---SerAspPheIleAsp 215
Db     482 TCCAGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGAC 541

Qy     216 LysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLys 235
Db     542 AGAAAGCTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGG 601

Qy     236 IleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerVal 255
Db     602 ATTATGAAGCTAAATTTAACTGAATGGCCTTATTTGTTGTTGGTGTATTTTGTGCCATT 661

Qy     256 LeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPhe 275
Db     662 ATAAATGGAGGCCTGCAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTT 721

Qy     276 GlyAsn---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 294
Db     722 ACAAGAATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTT 781

Qy     295 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 314
Db     782 CTAGCCCTTGGAATTATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAA 841

Qy     315 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 334
Db     842 GCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAG 901

Qy     335 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 354
Db     902 GATGTGAGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCC 961

Qy     355 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 374
Db     962 AATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAAT 1021

Qy     375 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 394
Db    1022 ATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTG 1081

Qy     395 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 414
Db    1082 TTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTG 1141

Qy     415 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 434
Db    1142 TCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAA 1201

Qy     435 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 454
Db    1202 GCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATG 1261

Qy     455 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 474
Db    1262 TATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGA 1321

Qy     475 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 494
Db    1322 ATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTT 1381

Qy     495 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 514
Db    1382 GGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCA 1441

Qy     515 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 534
Db    1442 GCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCC 1501

Qy     535 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 554
Db    1502 AAAGCCAAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGAC 1561

Qy     555 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 574
Db    1562 AGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAA 1621

Qy     575 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 594
Db    1622 GTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAG 1681

Qy     595 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 614
Db    1682 GTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTG 1741

Qy     615 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 634
Db    1742 GTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAA 1801

Qy     635 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 654
Db    1802 GAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAG 1861

Qy     655 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 674
Db    1862 CCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTG 1921

Qy     675 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 694
Db    1922 GTGTCACAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAG 1981

Qy     695 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 714
Db    1982 TCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGC 2041

Qy     715 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 734
Db    2042 CAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTG 2101

Qy     735 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 754
Db    2102 GATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGAC 2161

Qy     755 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 774
Db    2162 AAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAAT 2221

Qy     775 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 794
Db    2222 GCAGACTTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAG 2281

Qy     795 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 810
Db    2282 CTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 2329
```

RESULT 13
US-08-996-545-1
; Sequence 1, Application US/08996545
; Patent No. 5928898
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atrD of
; TITLE OF INVENTION: Aspergillus Nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:

```
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  U.S.
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/996,545
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Webster, Thomas D.
;      REGISTRATION NUMBER:  39,872
;      REFERENCE/DOCKET NUMBER:  X-11766
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-3334
;      TELEFAX:  317-276-2763
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4002 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..4002
US-08-996-545-1

Alignment Scores:
Pred. No.:              1.87e-187     Length:        4002
Score:                  1690.50       Matches:       366
Percent Similarity:     61.52%        Conservative:  152
Best Local Similarity:  43.47%        Mismatches:    293
Query Match:            41.44%        Indels:        31
DB:                     2             Gaps:          7

US-09-873-409-2 (1-812) x US-08-996-545-1 (1-4002)

Qy       1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
Db    1465 TTGCTGGATGGCCATGACATCAAGGACCTCAATCTCCGCTGGCTTCGCCAACAGATCTCT 1524

Qy      21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
Db    1525 TTGGTTAGCCAGGAGCCTGTTCTTTTGGCACGACGATTTATAAGAATATTAGGCACGGT 1584

Qy      41 --------------------------ArgAspAspValThrAspGluGluMetGluArg 51
Db    1585 CTCATCGGCACAAAGTACGAGAATGAATCCGAGGATAAGGTCCGGGAACTCATCGAGAAC 1644

Qy      52 AlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThr 71
Db    1645 GCGGCAAAAATGGCGAATGCTCATGACTTTATTACTGCCTTGCCTGAAGGTTATGAGACC 1704

Qy      72 LeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAla 91
Db    1705 AATGTTGGGCAGCGTGGCTTTCTCCTTTCAGGTGGCCAGAAACAGCGCATTGCAATCGCC 1764

Qy      92 ArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAsp 111
Db    1765 CGTGCCGTTGTTAGTGACCCAAAAATCCTGCTCCTGGATGAAGCTACTTCGGCCTTGGAC 1824

Qy     112 SerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThr 131
Db    1825 ACAAAATCCGAAGGCGTGGTTCAAGCAGCTTTGGAGAGGGCAGCTGAAGGCCGAACTACT 1884

Qy     132 IleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLys 151
Db    1885 ATTGTGATCGCTCATCGCCTTTCCACGATCAAAACGGCGCACAACATTGTGGTTCTGGTC 1944

Qy     152 AspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyr 171
Db    1945 AATGGCAAAATTGCTGAACAAGGAACTCACGATGAATTGGTTGACCGCGGAGGCGCTTAT 2004

Qy     172 TyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThr 191
Db    2005 CGCAAACTTGTGGAGGCTCAACGTATCAATGAACAGAAGGAAGCTGACGCCTTGGAGGAC 2064

Qy     192 TyrSerThrGluArgLysThrAsnSer------LeuProLeuHisSerValLysSerIle 209
Db    2065 GCCGACGCTGAGGATCTCACGAATGCAGATATTGCCAAAATCAAAACTGCGTCAAGCGCA 2124

Qy     210 LysSerAspPhe-----------------IleAspLys------AlaGluGluSerThr 221
Db    2125 TCATCCGATCTCGACGGAAAACCCACAACCATTGACCGCACGGGCACCCACAAGTCTGTT 2184

Qy     222 GlnSerLysGluIleSer-----------------LeuProGluValSerLeuLeuLys 235
Db    2185 TCCAGCGCGATTCTTTCTAAAAGACCCCCCGAAACAACTCCGAAATACTCATTATGGACG 2244

Qy     236 IleLeuLys------------LeuAsnLysProGluTrpProPheValValLeuGlyThr 251
Db    2245 CTGCTCAAATTTGTTGCTTCCTTCAACCGCCCTGAAATCCCGTACATGCTCATCGGTCTT 2304

Qy     252 LeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIle 271
Db    2305 GTCTTCTCAGTGTTAGCTGGTGGTGGCCAACCCACGCAAGCAGTGCTATATGCTAAAGCC 2364

Qy     272 IleThrMetPheGly------AsnAsnAspLysThrThrLeuLysHisAspAlaGluIle 289
Db    2365 ATCAGCACACTCTCGCTCCCAGAATCACAATATAGCAAGCTTCGACATGATGCGGATTTC 2424

Qy     290 TyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGly 309
Db    2425 TGGTCATTGATGTTCTTCGTGGTTGGTATCATTCAGTTTATCACGCAGTCAACCAATGGT 2484

Qy     310 LeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLys 329
Db    2485 GCTGCATTTGCCGTATGCTCCGAGAGACTTATTCGTCGCGCGAGAAGCACTGCCTTTCGG 2544

Qy     330 AlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeu 349
Db    2545 ACGATACTCCGTCAAGACATTGCTTTCTTTGACAAGGAAGAGAATAGCACCGGCGCTCTG 2604

Qy     350 ThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGly 369
Db    2605 ACCTCTTTCCTGTCCACCGAGACGAAGCATCTCTCCGGTGTTAGCGGTGTGACTCTAGGC 2664

Qy     370 ValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGly 389
Db    2665 ACGATCTTGATGACCTCCACGACCCTAGGAGCGGCTATCATTATTGCCCTGGCGATTGGG 2724

Qy     390 TrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIle 409
Db    2725 TGGAAATTGGCCTTAGTTTGTATCTCGGTTGTGCCGGTTCTCCTGGCATGCGGTTTCTAC 2784

Qy     410 GluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGly 429
Db    2785 CGATTCTATATGCTAGCCCAGTTTCAATCACGCTCCAAGCTTGCTTATGAGGGATCTGCA 2844

Qy     430 LysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLys 449
Db    2845 AACTTTGCTTGCGAGGCTACATCGTCTATCCGCACAGTTGCGTCATTAACCCGGGAAAGG 2904

Qy     450 AlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLys 469
Db    2905 GATGTCTGGGAGATTTACCATGCCCAGCTTGACGCACAAGGCAGGACCAGTCTAATCTCT 2964
```

```
Qy    470 AlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAla 489
                              |||||    |||   |||    ::::::|||       |||
Db   2965 GTCTTGAGGTCATCCCTGTTATATGCGTCGTCGCAGGCACTTGTTTTCTTCTGCGTTGCG 3024

Qy    490 AlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMet 509
             ||||||   :::|||      |||:::
Db   3025 CTCGGGTTTTGGTACGGAGGGACACTTCTTGGTCACCACGAGTATGACATTTTCCGCTTC 3084

Qy    510 PheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeu 529
          |||:::   |||:::   |||    :::||||||    :::    |||
Db   3085 TTTGTTTGTTTCTCCGAGATTCTCTTTGGTGCTCAATCCGCGGGCACCGTCTTTTCCTTT 3144

Qy    530 AlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLys 549
          ||||||:::         |||||||||:::   ||||||            |||    ::::::
Db   3145 GCACCAGACATGGGCAAGGCGAAGAATGCGGCCGCCGAATTCCGACGACTGTTCGACCGA 3204

Qy    550 LysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsn 569
          ||||||    ||||||:::   |||:::||||||:::|||    :::|||   ||||||
Db   3205 AAGCCACAAATTGATAACTGGTCTGAAGAGGGCGAGAAGCTCGAAACGGTGGAAGGTGAA 3264

Qy    570 LeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArg 589
          :::|||||||||   |||   |||   ||||||   ||||||:::         :::||||||
Db   3265 ATCGAATTTAGGAACGTGCACTTCAGATACCCGACCCGCCCAGAACAGCCTGTCCTGCGC 3324

Qy    590 GlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCys 609
          ||||||   |||:::::::::   |||:::   ||||||   ||||||   |||||||||
Db   3325 GGCTTGGACCTGACCGTGAAGCCTGGACAATATGTTGCGCTTGTCGGACCCAGCGGTTGT 3384

Qy    610 GlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnVal 629
          ||||||||||||::::::   ||||||:::|||   ||||||    :::   |||    :::
Db   3385 GGCAAGAGTACCACCATTGCATTGCTTGAGCGCTTTTACGATGCGATTGCCGGGTCCATC 3444

Qy    630 LeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAla 649
          |||   ||||||   |||          :::||||||:::          ||||||    ::::::
Db   3445 CTTGTTGATGGGAAGGACATAAGTAAACTAAATATCAACTCCTACCGCAGCTTTCTGTCA 3504

Qy    650 IleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGly 669
          :::|||   |||||||||   |||:::   :::|||   |||||||||         |||
Db   3505 CTGGTCAGCCAGGAGCCGACACTGTACCAGGGCACCATCAAGGAAAACATCTTACTTGGT 3564

Qy    670 AspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIle 689
                      ||||||   :::   :::   :::|||          |||||||||
Db   3565 ATTGTCGAAGATGACGTACCGGAAGAATTCTTGATTAAGGCTTGCAAGGACGCTAATATC 3624

Qy    690 HisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAla 709
          :::   ||||||         |||||||||   :::||||||   ||||||   ||||||
Db   3625 TACGACTTCATCATGTCGCTCCCGGAGGGCTTTAATACAGTTGTTGGCAGCAAGGGAGGC 3684

Qy    710 GlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysPro 729
             ||||||||||||||||||||||||:::||||||||||||||||||||:::    |||
Db   3685 ATGTTGTCTGGCGGCCAAAAGCAACGTGTGGCCATTGCCCGAGCCCTTCTTCGGGATCCC 3744

Qy    730 LysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValVal 749
          |||||||||||||||||||||||||||||||||||:::::::||||||||||||||||
Db   3745 AAAATCCTTCTTCTCGATGAAGCGACGTCAGCCCTCGACTCCGAGTCAGAAAAGGTCGTC 3804

Qy    750 GlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeu 769
          |||   |||||||||   |||         |||||||||    :::   |||   |||||||||
Db   3805 CAGGCGGCTTTGGATGCCGCTGCCCGAGGCCGAACCACAATCGCCGTTGCACACCGACTC 3864

Qy    770 SerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGln 789
          |||   ||||||   ||||||:::|||   |||         |||||||||   |||
Db   3865 AGCACGATTCAAAAGGCGGACGTTATCTATGTTTTCGACCAAGGCAAGATCGTCGAAAGC 3924

Qy    790 GlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGln 809
          |||||||||   ||||||::::::   :::      |||::::::|||||||||   |||
Db   3925 GGAACGCACAGCGAACTGGTCCAGAAAAAGGGCCGGTACTACGAGCTGGTCAACTTGCAG 3984

Qy    810 SerVal 811
          |||:::
Db   3985 AGCTTG 3990
```

RESULT 14

```
US-08-996-545-3
; Sequence 3, Application US/08996545
; Patent No. 5928898
;  GENERAL INFORMATION:
;    APPLICANT:  Skatrud, Paul L.
;    APPLICANT:  de Waard, Maarten A.
;    APPLICANT:  Peery, Robert B.
;    APPLICANT:  Andrade, Alan C.
;    TITLE OF INVENTION:  Multiple Drug Resistance Gene atrD of
;    TITLE OF INVENTION:  Aspergillus Nidulans
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  U.S.
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/996,545
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Webster, Thomas D.
;      REGISTRATION NUMBER:  39,872
;      REFERENCE/DOCKET NUMBER:  X-11766
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-3334
;      TELEFAX:  317-276-2763
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4002 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  mRNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-996-545-3

Alignment Scores:
Pred. No.:               1.87e-187      Length:         4002
Score:                   1690.50        Matches:        366
Percent Similarity:      61.52%         Conservative:   152
Best Local Similarity:   43.47%         Mismatches:     293
Query Match:             41.44%         Indels:         31
DB:                      2              Gaps:           7

US-09-873-409-2 (1-812) x US-08-996-545-3 (1-4002)

Qy       1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
            ::::::|||   :::||||||:::      ||||||:::|||        |||      |||
Db    1465 UUGCUGGAUGGCCAUGACAUCAAGGACCUCAAUCUCCGCUGGCUUCGCCAACAGAUCUCU 1524

Qy      21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
           :::|||||||||||||||||||||||||||||||||||         ||||||::::::|||
Db    1525 UUGGUUAGCCAGGAGCCUGUUCUUUUUGGCACGACGAUUUAUAAGAAUAUUAGGCACGGU 1584

Qy      41 ---------------------------ArgAspAspValThrAspGluGluMetGluArg 51
                                      :::|||         |||   :::|||
Db    1585 CUCAUCGGCACAAAGUACGAGAAUGAAUCCGAGGAUAAGGUCCGGGAACUCAUCGAGAAC 1644

Qy      52 AlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThr 71
           ||||||:::   |||||||||:::||||||||         |||      :::   |||
Db    1645 GCGGCAAAAAUGGCGAAUGCUCAUGACUUUAUUACUGCCUUGCCUGAAGGUUAUGAGACC 1704

Qy      72 LeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAla 91
              ||||||::::::|||      :::|||||||||||||||||||||||||||||
```

```
Db   1705 AAUGUUGGGCAGCGUGGCUUUCUCCUUUCAGGUGGCCAGAAACAGCGCAUUGCAAUCGCC 1764

Qy     92 ArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAsp 111
Db   1765 CGUGCCGUUGUUAGUGACCCAAAAAUCCUGCUCCUGGAUGAAGCUACUUCGGCCUUGGAC 1824

Qy    112 SerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThr 131
Db   1825 ACAAAAUCCGAAGGCGUGGUUCAAGCAGCUUUGGAGAGGGCAGCUGAAGGCCGAACUACU 1884

Qy    132 IleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLys 151
Db   1885 AUUGUGAUCGCUCAUCGCCUUUCCACGAUCAAAACGGCGCACAACAUUGUGGUUCUGGUC 1944

Qy    152 AspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyr 171
Db   1945 AAUGGCAAAAUUGCUGAACAAGGAACUCACGAUGAAUUGGUUGACCGCGGAGGCGCUUAU 2004

Qy    172 TyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThr 191
Db   2005 CGCAAACUUGUGGAGGCUCAACGUAUCAAUGAACAGAAGGAAGCUGACGCCUUGGAGGAC 2064

Qy    192 TyrSerThrGluArgLysThrAsnSer------LeuProLeuHisSerValLysSerIle 209
Db   2065 GCCGACGCUGAGGAUCUCACGAAUGCAGAUAUUGCCAAAAUCAAAACUGCGUCAAGCGCA 2124

Qy    210 LysSerAspPhe------------------IleAspLys------AlaGluGluSerThr 221
Db   2125 UCAUCCGAUCUCGACGGAAAACCCACAACCAUUGACCGCACGGGCACCCACAAGUCUGUU 2184

Qy    222 GlnSerLysGluIleSer------------------LeuProGluValSerLeuLeuLys 235
Db   2185 UCCAGCGCGAUUCUUUCUAAAAGACCCCCCGAAACAACUCCGAAAUACUCAUUAUGGACG 2244

Qy    236 IleLeuLys------------LeuAsnLysProGluTrpProPheValValLeuGlyThr 251
Db   2245 CUGCUCAAAUUUGUUGCUUCCUUCAACCGCCCUGAAAUCCCGUACAUGCUCAUCGGUCUU 2304

Qy    252 LeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIle 271
Db   2305 GUCUUCUCAGUGUUAGCUGGUGGUGGCCAACCCACGCAAGCAGUGCUAUAUGCUAAAGCC 2364

Qy    272 IleThrMetPheGly------AsnAsnAspLysThrThrLeuLysHisAspAlaGluIle 289
Db   2365 AUCAGCACACUCUCGCUCCCAGAAUCACAAUAUAGCAAGCUUCGACAUGAUGCGGAUUUC 2424

Qy    290 TyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGly 309
Db   2425 UGGUCAUUGAUGUUCUUCGUGGUUGGUAUCAUUCAGUUUAUCACGCAGUCAACCAAUGGU 2484

Qy    310 LeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLys 329
Db   2485 GCUGCAUUUGCCGUAUGCUCCGAGAGACUUAUUCGUCGCGCGAGAAGCACUGCCUUUCGG 2544

Qy    330 AlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeu 349
Db   2545 ACGAUACUCCGUCAAGACAUUGCUUUCUUUGACAAGGAAGAGAAUAGCACCGGCGCUCUG 2604

Qy    350 ThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGly 369
Db   2605 ACCUCUUUCCUGUCCACCGAGACGAAGCAUCUCUCCGGUGUUAGCGGUGUGACUCUAGGC 2664

Qy    370 ValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGly 389
Db   2665 ACGAUCUUGAUGACCUCCACGACCCUAGGAGCGGCUAUCAUUAUUGCCCUGGCGAUUGGG 2724

Qy    390 TrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIle 409
Db   2725 UGGAAAUUGGCCUUAGUUUGUAUCUCGGUUGUGCCGGUUCUCCUGGCAUGCGGUUUCUAC 2784

Qy    410 GluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGly 429
Db   2785 CGAUUCUAUAUGCUAGCCCAGUUUCAAUCACGCUCCAAGCUUGCUUAUGAGGGAUCUGCA 2844

Qy    430 LysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLys 449
Db   2845 AACUUUGCUUGCGAGGCUACAUCGUCUAUCCGCACAGUUGCGUCAUUAACCCGGGAAAGG 2904

Qy    450 AlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLys 469
Db   2905 GAUGUCUGGGAGAUUUACCAUGCCCAGCUUGACGCACAAGGCAGGACCAGUCUAAUCUCU 2964

Qy    470 AlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAla 489
Db   2965 GUCUUGAGGUCAUCCCUGUUAUAUGCGUCGUCGCAGGCACUUGUUUUCUUCUGCGUUGCG 3024

Qy    490 AlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMet 509
Db   3025 CUCGGGUUUUGGUACGGAGGGACACUUCUUGGUCACCACGAGUAUGACAUUUUCCGCUUC 3084

Qy    510 PheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeu 529
Db   3085 UUUGUUUGUUUCUCCGAGAUUCUCUUUGGUGCUCAAUCCGCGGGCACCGUCUUUUCCUUU 3144

Qy    530 AlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLys 549
Db   3145 GCACCAGACAUGGGCAAGGCGAAGAAUGCGGCCGCCGAAUUCCGACGACUGUUCGACCGA 3204

Qy    550 LysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsn 569
Db   3205 AAGCCACAAAUUGAUAACUGGUCUGAAGAGGGCGAGAAGCUCGAAACGGUGGAAGGUGAA 3264

Qy    570 LeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArg 589
Db   3265 AUCGAAUUUAGGAACGUGCACUUCAGAUACCCGACCCGCCCAGAACAGCCUGUCCUGCGC 3324

Qy    590 GlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCys 609
Db   3325 GGCUUGGACCUGACCGUGAAGCCUGGACAAUAUGUUGCGCUUGUCGGACCCAGCGGUUGU 3384

Qy    610 GlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnVal 629
Db   3385 GGCAAGAGUACCACCAUUGCAUUGCUUGAGCGCUUUUACGAUGCGAUUGCCGGGUCCAUC 3444

Qy    630 LeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAla 649
Db   3445 CUUGUUGAUGGGAAGGACAUAAGUAAACUAAAUAUCAACUCCUACCGCAGCUUUCUGUCA 3504

Qy    650 IleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGly 669
Db   3505 CUGGUCAGCCAGGAGCCGACACUGUACCAGGGCACCAUCAAGGAAAACAUCUUACUUGGU 3564

Qy    670 AspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIle 689
Db   3565 AUUGUCGAAGAUGACGUACCGGAAGAAUUCUUGAUUAAGGCUUGCAAGGACGCUAAUAUC 3624

Qy    690 HisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAla 709
Db   3625 UACGACUUCAUCAUGUCGCUCCCGGAGGGCUUUAAUACAGUUGUUGGCAGCAAGGGAGGC 3684

Qy    710 GlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysPro 729
Db   3685 AUGUUGUCUGGCGGCCAAAAGCAACGUGUGGCCAUUGCCCGAGCCCUUCUUCGGGAUCCC 3744

Qy    730 LysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValVal 749
Db   3745 AAAAUCCUUCUUCUCGAUGAAGCGACGUCAGCCCUCGACUCCGAGUCAGAAAAGGUCGUC 3804

Qy    750 GlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeu 769
Db   3805 CAGGCGGCUUUGGAUGCCGCUGCCCGAGGCCGAACCACAAUCGCCGUUGCACACCGACUC 3864

Qy    770 SerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGln 789
Db   3865 AGCACGAUUCAAAAGGCGGACGUUAUCUAUGUUUUCGACCAAGGCAAGAUCGUCGAAAGC 3924
```

```
Qy     790 GlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGln 809
           |||||||||  ||||||:::::  :::       |||::::::||||||||    |||
Db    3925 GGAACGCACAGCGAACUGGUCCAGAAAAAGGGCCGGUACUACGAGCUGGUCAACUUGCAG 3984

Qy     810 SerVal 811
           |||:::
Db    3985 AGCUUG 3990

RESULT 15
US-09-328-320-1
; Sequence 1, Application US/09328320
; Patent No. 6228615
;  GENERAL INFORMATION:
;    APPLICANT:  Skatrud, Paul L.
;    APPLICANT:  de Waard, Maarten A.
;    APPLICANT:  Peery, Robert B.
;    APPLICANT:  Andrade, Alan C.
;    TITLE OF INVENTION:  Multiple Drug Resistance Gene atrD of
;    TITLE OF INVENTION:  Aspergillus Nidulans
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  U.S.
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/328,320
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/996,545
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Webster, Thomas D.
;      REGISTRATION NUMBER:  39,872
;      REFERENCE/DOCKET NUMBER:  X-11766
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-3334
;      TELEFAX:  317-276-2763
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4002 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..4002
US-09-328-320-1

Alignment Scores:
Pred. No.:              1.87e-187     Length:          4002
Score:                  1690.50       Matches:         366
Percent Similarity:     61.52%        Conservative:    152
Best Local Similarity:  43.47%        Mismatches:      293
Query Match:            41.44%        Indels:          31
DB:                     4             Gaps:            7

US-09-873-409-2 (1-812) x US-09-328-320-1 (1-4002)

Qy       1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
           ::::::|||   :::||||||:::    ||||||:::|||       |||      |||
Db    1465 TTGCTGGATGGCCATGACATCAAGGACCTCAATCTCCGCTGGCTTCGCCAACAGATCTCT 1524
```

```
Qy      21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
           :::|||||||||||||||||||||||||||||||||||         ||||||::::::|||
Db    1525 TTGGTTAGCCAGGAGCCTGTTCTTTTTGGCACGACGATTTATAAGAATATTAGGCACGGT 1584

Qy      41 --------------------------ArgAspAspValThrAspGluGluMetGluArg 51
                                     :::|||          |||   :::|||
Db    1585 CTCATCGGCACAAAGTACGAGAATGAATCCGAGGATAAGGTCCGGGAACTCATCGAGAAC 1644

Qy      52 AlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThr 71
           ||||||:::   |||||||||:::||||||||          |||       :::   |||
Db    1645 GCGGCAAAAATGGCGAATGCTCATGACTTTATTACTGCCTTGCCTGAAGGTTATGAGACC 1704

Qy      72 LeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAla 91
              ||||||::::::|||      :::||||||||||||||||||||||||||||||||||
Db    1705 AATGTTGGGCAGCGTGGCTTTCTCCTTTCAGGTGGCCAGAAACAGCGCATTGCAATCGCC 1764

Qy      92 ArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAsp 111
           |||||||:::|||   :::||||||||||:::||||||||||||||||||||||||||||
Db    1765 CGTGCCGTTGTTAGTGACCCAAAAATCCTGCTCCTGGATGAAGCTACTTCGGCCTTGGAC 1824

Qy     112 SerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThr 131
           ::::::|||:::      ||||||||||||||||||:::|||::::::||||||||||||
Db    1825 ACAAAATCCGAAGGCGTGGTTCAAGCAGCTTTGGAGAGGGCAGCTGAAGGCCGAACTACT 1884

Qy     132 IleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLys 151
           ||||||:::||||||||||||||||||||||||::::::|||       ||||||   |||
Db    1885 ATTGTGATCGCTCATCGCCTTTCCACGATCAAAACGGCGCACAACATTGTGGTTCTGGTC 1944

Qy     152 AspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyr 171
           :::|||   :::||||||:::|||   |||   ||||||:::   :::   |||   |||
Db    1945 AATGGCAAAATTGCTGAACAAGGAACTCACGATGAATTGGTTGACCGCGGAGGCGCTTAT 2004

Qy     172 TyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThr 191
                 ||||||   :::|||   |||   :::      |||
Db    2005 CGCAAACTTGTGGAGGCTCAACGTATCAATGAACAGAAGGAAGCTGACGCCTTGGAGGAC 2064

Qy     192 TyrSerThrGluArgLysThrAsnSer------LeuProLeuHisSerValLysSerIle 209
                    |||   ||||||:::            :::   :::      |||
Db    2065 GCCGACGCTGAGGATCTCACGAATGCAGATATTGCCAAAATCAAAACTGCGTCAAGCGCA 2124

Qy     210 LysSerAspPhe------------------IleAspLys------AlaGluGluSerThr 221
              |||||                      ||||||:::            :::|||
Db    2125 TCATCCGATCTCGACGGAAAACCCACAACCATTGACCGCACGGGCACCCACAAGTCTGTT 2184

Qy     222 GlnSerLysGluIleSer------------------LeuProGluValSerLeuLeuLys 235
              |||      :::|||                     |||:::   ||||||
Db    2185 TCCAGCGCGATTCTTTCTAAAAGACCCCCGAAACAACTCCGAAATACTCATTATGGACG 2244

Qy     236 IleLeuLys------------LeuAsnLysProGluTrpProPheValValLeuGlyThr 251
           :::||||||             |||:::||||||   |||:::::::::::::|||
Db    2245 CTGCTCAAATTTGTTGCTTCCTTCAACCGCCCTGAAATCCCGTACATGCTCATCGGTCTT 2304

Qy     252 LeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIle 271
           :::   |||||||||   |||          |||      ::::::::::::||||||
Db    2305 GTCTTCTCAGTGTTAGCTGGTGGTGGCCAACCCACGCAAGCAGTGCTATATGCTAAAGCC 2364

Qy     272 IleThrMetPheGly------AsnAsnAspLysThrThrLeuLysHisAspAlaGluIle 289
           |||:::                :::      :::   |||:::|||||||||:::
Db    2365 ATCAGCACACTCTCGCTCCCAGAATCACAATATAGCAAGCTTCGACATGATGCGGATTTC 2424

Qy     290 TyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGly 309
           :::||||::::::|||   ::::::|||:::|||   |||::::::                |||
Db    2425 TGGTCATTGATGTTCTTCGTGGTTGGTATCATTCAGTTTATCACGCAGTCAACCAATGGT 2484

Qy     310 LeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLys 329
                 :::         |||   |||         |||   |||      ||||||:::
Db    2485 GCTGCATTTGCCGTATGCTCCGAGAGACTTATTCGTCGCGCGAGAAGCACTGCCTTTCGG 2544

Qy     330 AlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeu 349
              :::|||   ||||||||||||:::||||||::::::||||||||||||||    |||
Db    2545 ACGATACTCCGTCAAGACATTGCTTTCTTTGACAAGGAAGAGAATAGCACCGGCGCTCTG 2604
```

```
Qy     350 ThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGly 369
           |||:::   |||:::   :::         :::   |||   :::|||      :::|||
Db    2605 ACCTCTTTCCTGTCCACCGAGACGAAGCATCTCTCCGGTGTTAGCGGTGTGACTCTAGGC 2664

Qy     370 ValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGly 389
           :::               :::|||   :::|||   ::::::||||||:::      |||
Db    2665 ACGATCTTGATGACCTCCACGACCCTAGGAGCGGCTATCATTATTGCCCTGGCGATTGGG 2724

Qy     390 TrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIle 409
           |||::::::      :::   :::|||:::   ||||||||            |||
Db    2725 TGGAAATTGGCCTTAGTTTGTATCTCGGTTGTGCCGGTTCTCCTGGCATGCGGTTTCTAC 2784

Qy     410 GluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGly 429
                             :::      |||   ::::::   |||         :::   :::
Db    2785 CGATTCTATATGCTAGCCCAGTTTCAATCACGCTCCAAGCTTGCTTATGAGGGATCTGCA 2844

Qy     430 LysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLys 449
                    |||   ||||||         :::|||||||||:::   ||||||||||||||:::
Db    2845 AACTTTGCTTGCGAGGCTACATCGTCTATCCGCACAGTTGCGTCATTAACCCGGGAAAGG 2904

Qy     450 AlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLys 469
                 ::::::|||            |||         |||   |||   :::
Db    2905 GATGTCTGGGAGATTTACCATGCCCAGCTTGACGCACAAGGCAGGACCAGTCTAATCTCT 2964

Qy     470 AlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAla 489
                                 ||||||   |||   |||   ::::::|||      |||
Db    2965 GTCTTGAGGTCATCCCTGTTATATGCGTCGTCGCAGGCACTTGTTTCTTCTGCGTTGCG 3024

Qy     490 AlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMet 509
              ||||||   :::|||   ||||:::
Db    3025 CTCGGGTTTTGGTACGGAGGGACACTTCTTGGTCACCACGAGTATGACATTTTCCGCTTC 3084

Qy     510 PheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeu 529
           |||:::   |||:::   |||   :::||||||   :::   |||
Db    3085 TTTGTTTGTTTCTCCGAGATTCTCTTTGGTGCTCAATCCGCGGGCACCGTCTTTTCCTTT 3144

Qy     530 AlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLys 549
           ||||||:::      |||||||||:::   ||||||            |||   ::::::
Db    3145 GCACCAGACATGGGCAAGGCGAAGAATGCGGCCGCCGAATTCCGACGACTGTTCGACCGA 3204

Qy     550 LysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsn 569
           ||||||   ||||||:::   |||:::||||||:::|||   :::|||   ||||||
Db    3205 AAGCCACAAATTGATAACTGGTCTGAAGAGGGCGAGAAGCTCGAAACGGTGGAAGGTGAA 3264

Qy     570 LeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArg 589
           :::|||||||||   |||   |||   ||||||   ||||||:::      :::||||||
Db    3265 ATCGAATTTAGGAACGTGCACTTCAGATACCCGACCCGCCCAGAACAGCCTGTCCTGCGC 3324

Qy     590 GlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCys 609
           ||||||   |||:::::::::   |||:::   ||||||   ||||||   |||||||||
Db    3325 GGCTTGGACCTGACCGTGAAGCCTGGACAATATGTTGCGCTTGTCGGACCCAGCGGTTGT 3384

Qy     610 GlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnVal 629
           ||||||||||||::::::   ||||||:::|||   ||||||   :::   |||   :::
Db    3385 GGCAAGAGTACCACCATTGCATTGCTTGAGCGCTTTTACGATGCGATTGCCGGGTCCATC 3444

Qy     630 LeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAla 649
           |||   ||||||   |||         :::||||||:::         ||||||   ::::::
Db    3445 CTTGTTGATGGGAAGGACATAAGTAAACTAAATATCAACTCCTACCGCAGCTTTCTGTCA 3504

Qy     650 IleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGly 669
           :::|||   |||||||||   |||:::         :::|||   |||||||||      |||
Db    3505 CTGGTCAGCCAGGAGCCGACACTGTACCAGGGCACCATCAAGGAAAACATCTTACTTGGT 3564

Qy     670 AspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIle 689
                             ||||||   :::   :::   :::|||         |||||||||
Db    3565 ATTGTCGAAGATGACGTACCGGAAGAATTCTTGATTAAGGCTTGCAAGGACGCTAATATC 3624

Qy     690 HisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAla 709
           :::   ||||||      |||||||||   :::||||||         ||||||
Db    3625 TACGACTTCATCATGTCGCTCCCGGAGGGCTTTAATACAGTTGTTGGCAGCAAGGGAGGC 3684

Qy     710 GlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysPro 729
           ||||||||||||||||||||||:::||||||||||||||||||:::   |||
Db    3685 ATGTTGTCTGGCGGCCAAAAGCAACGTGTGGCCATTGCCCGAGCCCTTCTTCGGGATCCC 3744

Qy     730 LysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValVal 749
           ||||||||||||||||||||||||||||||||||||::::::||||||||||||||||||
Db    3745 AAAATCCTTCTTCTCGATGAAGCGACGTCAGCCCTCGACTCCGAGTCAGAAAAGGTCGTC 3804

Qy     750 GlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeu 769
           |||   |||||||||   |||      |||||||||   :::   |||   |||||||||
Db    3805 CAGGCGGCTTTGGATGCCGCTGCCCGAGGCCGAACCACAATCGCCGTTGCACACCGACTC 3864

Qy     770 SerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGln 789
           |||   ||||||   ||||||:::|||   |||         |||||||||   |||
Db    3865 AGCACGATTCAAAAGGCGGACGTTATCTATGTTTTCGACCAAGGCAAGATCGTCGAAAGC 3924

Qy     790 GlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGln 809
           |||||||||   ||||||::::::   :::      |||:::::::||||||||   |||
Db    3925 GGAACGCACAGCGAACTGGTCCAGAAAAAGGGCCGGTACTACGAGCTGGTCAACTTGCAG 3984

Qy     810 SerVal 811
           |||:::
Db    3985 AGCTTG 3990
```

Search completed: March 31, 2003, 15:23:06
Job time : 266.913 secs

```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 31, 2003, 13:52:37 ; Search time 295.562 Seconds
                                           (without alignments)
                                           2337.888 Million cell updates/sec

Title:          US-09-873-409-2
Perfect score:  4079
Sequence:       1 MVDENDIRALNVRHYRDHIG..........QELLRNRDIYFKLVNAQSVQ 812

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters:      1148742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115424_19290/app_query.fasta_1.7544
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09873409_@CGN_1_1_960_@runat_27032003_115424_19290
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Published_Applications_NA:*
                1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
                2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
                3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
                4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
                5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
                6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
                7:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
                8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
                9:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
                10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
                11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
                12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
                13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
                14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.
```

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	4079	100.0	2856	10	US-09-873-409-10	Sequence 10, Appl
	2	4079	100.0	3177	10	US-09-873-409-12	Sequence 12, Appl
	3	4079	100.0	3621	10	US-09-873-409-14	Sequence 14, Appl
	4	4079	100.0	3702	10	US-09-873-409-13	Sequence 13, Appl
	5	3456	84.7	2066	10	US-09-873-409-9	Sequence 9, Appli
	6	2465.5	60.4	3924	10	US-09-880-107-2299	Sequence 2299, Ap
	7	2441	59.8	4369	10	US-09-769-097-1	Sequence 1, Appli
	8	2441	59.8	4425	10	US-09-769-097-3	Sequence 3, Appli
	9	2438	59.8	4317	9	US-10-044-671-1	Sequence 1, Appli
	10	2436	59.7	3860	10	US-09-866-866A-1	Sequence 1, Appli
	11	2436	59.7	3860	10	US-09-866-866A-3	Sequence 3, Appli
	12	2436	59.7	4643	9	US-10-072-621-2	Sequence 2, Appli
	13	2436	59.7	8630	10	US-09-306-417-1	Sequence 1, Appli
	14	2436	59.7	8630	10	US-09-306-417-2	Sequence 2, Appli
	15	2431.5	59.6	3912	10	US-09-917-800A-1560	Sequence 1560, Ap
	16	2425	59.5	4788	10	US-09-866-866A-7	Sequence 7, Appli
	17	2404.5	58.9	4189	10	US-09-866-866A-5	Sequence 5, Appli
	18	2372	58.2	4254	10	US-09-917-800A-1424	Sequence 1424, Ap
	19	2013.5	49.4	5010	10	US-09-917-800A-483	Sequence 483, App
	20	1694	41.5	4175	10	US-09-749-340-3	Sequence 3, Appli
	21	1690.5	41.4	4002	9	US-09-758-828-1	Sequence 1, Appli
	22	1690.5	41.4	4002	9	US-09-758-828-3	Sequence 3, Appli
	23	1678.5	41.1	3512	10	US-09-749-340-7	Sequence 7, Appli
	24	1669.5	40.9	3861	9	US-09-938-842A-2237	Sequence 2237, Ap
	25	1617	39.6	2698	10	US-09-749-340-5	Sequence 5, Appli
	26	1550.5	38.0	3861	9	US-09-938-842A-263	Sequence 263, App
	27	1538.5	37.7	2681	10	US-09-749-340-8	Sequence 8, Appli
	28	1452.5	35.6	4653	9	US-10-101-388-2	Sequence 2, Appli
	29	1287.5	31.6	8036	9	US-10-101-388-1	Sequence 1, Appli
	30	1116.5	27.4	3792	9	US-09-882-694-10	Sequence 10, Appl
	31	1032.5	25.3	3999	9	US-09-882-694-9	Sequence 9, Appli
	32	866	21.2	7296	10	US-09-070-927A-59	Sequence 59, Appl
c	33	864	21.2	9047	10	US-09-070-927A-12	Sequence 12, Appl
	34	836.5	20.5	1810	10	US-09-749-340-4	Sequence 4, Appli
	35	828	20.3	1940	10	US-09-873-409-16	Sequence 16, Appl
	36	820.5	20.1	3741	9	US-09-738-626-1251	Sequence 1251, Ap
	37	812.5	19.9	2021	10	US-09-873-409-15	Sequence 15, Appl
c	38	807	19.8	3601	10	US-09-070-927A-255	Sequence 255, App
	39	788	19.3	12438	10	US-09-070-927A-173	Sequence 173, App
c	40	770	18.9	6415	10	US-09-070-927A-275	Sequence 275, App
	41	769	18.9	640681	10	US-09-790-988-1	Sequence 1, Appli
	42	758.5	18.6	2298	9	US-10-156-239-18	Sequence 18, Appl
	43	758.5	18.6	2298	10	US-09-795-693-18	Sequence 18, Appl
	44	758.5	18.6	3408	9	US-10-156-239-16	Sequence 16, Appl
	45	758.5	18.6	3408	10	US-09-795-693-16	Sequence 16, Appl

ALIGNMENTS

```
RESULT 1
US-09-873-409-10
; Sequence 10, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
;   LENGTH: 2856
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-409-10

Alignment Scores:
Pred. No.:              0          Length:        2856
Score:                  4079.00    Matches:       812
Percent Similarity:     100.00%    Conservative:  0
Best Local Similarity:  100.00%    Mismatches:    0
Query Match:            100.00%    Indels:        0
DB:                     10         Gaps:          0
```

```
US-09-873-409-2 (1-812) x US-09-873-409-10 (1-2856)

Qy       1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     418 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 477

Qy      21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     478 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 537

Qy      41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     538 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 597

Qy      61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     598 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 657

Qy      81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     658 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 717

Qy     101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     718 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 777

Qy     121 AlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThr 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     778 GCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTAGCACACCGACTTTCTACT 837

Qy     141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     838 ATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATGCTGGCGGAGAAAGGAGCA 897

Qy     161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     898 CATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTTGTGATGTCACAGGATATT 957

Qy     181 LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     958 AAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCA 1017

Qy     201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSer 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1018 CTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCC 1077

Qy     221 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1078 ACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAAC 1137

Qy     241 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1138 AAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCTGTTCTAAATGGAACTGTT 1197

Qy     261 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1198 CATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATGTTTGGAAATAATGATAAA 1257

Qy     281 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1258 ACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTCGTCATTTTGGGTGTTATT 1317

Qy     301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1318 TGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGAGCAGGGGAAATTTTAACG 1377

Qy     321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1378 ATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAGGATATTGCCTGGTTTGAT 1437

Qy     341 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1438 GAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCCATAGATATAGCACAAATT 1497

Qy     361 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1498 CAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAATGCAACTAACATGGGACTT 1557

Qy     381 SerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAla 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1558 TCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTCCTGATTCTGAGTATTGCT 1617

Qy     401 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1618 CCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATGACTGGATTTGCCAACAAA 1677

Qy     421 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1678 GATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAGAATATACGT 1737

Qy     441 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1738 ACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATGTATGAAGAGATGCTTCAG 1797

Qy     461 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1798 ACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTATGCATTCAGC 1857

Qy     481 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1858 CATGCCTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTTGGAGCCTATTTAATTCAA 1917

Qy     501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1918 GCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACTGCAATTGCATATGGAGCT 1977

Qy     521 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1978 ATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCCAAAGCCAAATCGGGGGCT 2037

Qy     541 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2038 GCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGACAGCCGCAGTCAAGAAGGG 2097

Qy     561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2098 AAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTCTTCTATCCA 2157

Qy     581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2158 TGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACA 2217

Qy     601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2218 GTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCTGTTCAACTTCTGCAGAGA 2277

Qy     621 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2278 CTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTGGATGCAAAAGAATTGAAT 2337

Qy     641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2338 GTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAGCCTGTGCTCTTCAACTGC 2397

Qy     661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2398 AGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTGGTGCCATTAGATGAGATC 2457

Qy     681 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2458 AAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAAGGTCTCCCTGAGAAATAC 2517

Qy     701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2518 AACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCT 2577
```

```
Qy     721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2578 ATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCC 2637

Qy     741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2638 CTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGATAAAGCCAGGACGGGAAGG 2697

Qy     761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2698 ACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGGTT 2757

Qy     781 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2758 CTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAGCTCCTGAGAAATCGAGAC 2817

Qy     801 IleTyrPheLysLeuValAsnAlaGlnSerValGln 812
           ||||||||||||||||||||||||||||||||||||
Db    2818 ATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 2853
```

RESULT 2
US-09-873-409-12
; Sequence 12, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
;   LENGTH: 3177
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: No. US20020037522A1e
;   LOCATION: (198)..(198)
;   OTHER INFORMATION: n at position 198 represents any nucleotide (A, T, C or G)
US-09-873-409-12

Alignment Scores:			
Pred. No.:	0	Length:	3177
Score:	4079.00	Matches:	812
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-873-409-2 (1-812) x US-09-873-409-12 (1-3177)

```
Qy       1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     739 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 798

Qy      21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     799 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 858

Qy      41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     859 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 918

Qy      61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     919 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 978

Qy      81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     979 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 1038

Qy     101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1039 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 1098

Qy     121 AlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThr 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1099 GCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTAGCACACCGACTTTCTACT 1158

Qy     141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1159 ATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATGCTGGCGGAGAAAGGAGCA 1218

Qy     161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1219 CATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTTGTGATGTCACAGGATATT 1278

Qy     181 LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1279 AAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCA 1338

Qy     201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSer 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1339 CTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCC 1398

Qy     221 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1399 ACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAAC 1458

Qy     241 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1459 AAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCTGTTCTAAATGGAACTGTT 1518

Qy     261 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1519 CATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATGTTTGGAAATAATGATAAA 1578

Qy     281 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1579 ACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTCGTCATTTTGGGTGTTATT 1638

Qy     301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1639 TGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGAGCAGGGGAAATTTTAACG 1698

Qy     321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1699 ATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAGGATATTGCCTGGTTTGAT 1758

Qy     341 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1759 GAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCCATAGATATAGCACAAATT 1818

Qy     361 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1819 CAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAATGCAACTAACATGGGACTT 1878

Qy     381 SerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAla 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1879 TCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTCCTGATTCTGAGTATTGCT 1938

Qy     401 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1939 CCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATGACTGGATTTGCCAACAAA 1998

Qy     421 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1999 GATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAGAATATACGT 2058

Qy     441 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2059 ACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATGTATGAAGAGATGCTTCAG 2118
```

```
Qy    461 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2119 ACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTATGCATTCAGC 2178

Qy    481 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2179 CATGCCTTTATATATTTGCCTATGCAGCAGGGTTTCGATTTGGAGCCTATTTAATTCAA 2238

Qy    501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2239 GCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACTGCAATTGCATATGGAGCT 2298

Qy    521 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2299 ATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCCAAAGCCAAATCGGGGGCT 2358

Qy    541 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2359 GCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGACAGCCGCAGTCAAGAAGGG 2418

Qy    561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2419 AAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTCTTCTATCCA 2478

Qy    581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2479 TGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACA 2538

Qy    601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2539 GTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCTGTTCAACTTCTGCAGAGA 2598

Qy    621 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2599 CTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTGGATGCAAAAGAATTGAAT 2658

Qy    641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2659 GTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAGCCTGTGCTCTTCAACTGC 2718

Qy    661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2719 AGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTGGTGCCATTAGATGAGATC 2778

Qy    681 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2779 AAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAAGGTCTCCCTGAGAAATAC 2838

Qy    701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2839 AACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCT 2898

Qy    721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2899 ATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCC 2958

Qy    741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2959 CTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGATAAAGCCAGGACGGGAAGG 3018

Qy    761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3019 ACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGGTT 3078

Qy    781 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3079 CTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAGCTCCTGAGAAATCGAGAC 3138

Qy    801 IleTyrPheLysLeuValAsnAlaGlnSerValGln 812
          ||||||||||||||||||||||||||||||||||||
Db   3139 ATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3174

RESULT 3
US-09-873-409-14
; Sequence 14, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
;   LENGTH: 3621
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-409-14

Alignment Scores:
Pred. No.:              0              Length:        3621
Score:                  4079.00        Matches:       812
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     10             Gaps:          0

US-09-873-409-2 (1-812) x US-09-873-409-14 (1-3621)

Qy      1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1183 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 1242

Qy     21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1243 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 1302

Qy     41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1303 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 1362

Qy     61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1363 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 1422

Qy     81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1423 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 1482

Qy    101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1483 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 1542

Qy    121 AlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThr 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1543 GCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTAGCACACCGACTTTCTACT 1602

Qy    141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1603 ATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATGCTGGCGGAGAAAGGAGCA 1662

Qy    161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1663 CATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTTGTGATGTCACAGGATATT 1722

Qy    181 LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1723 AAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCA 1782

Qy    201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSer 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1783 CTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCC 1842
```

```
Qy      221 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1843 ACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAAC 1902

Qy      241 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1903 AAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCTGTTCTAAATGGAACTGTT 1962

Qy      261 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1963 CATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATGTTTGGAAATAATGATAAA 2022

Qy      281 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2023 ACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTCGTCATTTTGGGTGTTATT 2082

Qy      301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2083 TGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGAGCAGGGGAAATTTTAACG 2142

Qy      321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2143 ATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAGGATATTGCCTGGTTTGAT 2202

Qy      341 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2203 GAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCCATAGATATAGCACAAATT 2262

Qy      361 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2263 CAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAATGCAACTAACATGGGACTT 2322

Qy      381 SerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAla 400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2323 TCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTCCTGATTCTGAGTATTGCT 2382

Qy      401 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2383 CCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATGACTGGATTTGCCAACAAA 2442

Qy      421 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2443 GATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAGAATATACGT 2502

Qy      441 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2503 ACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATGTATGAAGAGATGCTTCAG 2562

Qy      461 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2563 ACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTATGCATTCAGC 2622

Qy      481 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2623 CATGCCTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTTGGAGCCTATTTAATTCAA 2682

Qy      501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2683 GCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACTGCAATTGCATATGGAGCT 2742

Qy      521 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2743 ATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCCAAAGCCAAATCGGGGGCT 2802

Qy      541 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2803 GCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGACAGCCGCAGTCAAGAAGGG 2862

Qy      561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2863 AAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTCTTCTATCCA 2922

Qy      581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2923 TGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACA 2982

Qy      601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2983 GTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCTGTTCAACTTCTGCAGAGA 3042

Qy      621 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3043 CTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTGGATGCAAAAGAATTGAAT 3102

Qy      641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3103 GTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAGCCTGTGCTCTTCAACTGC 3162

Qy      661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3163 AGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTGGTGCCATTAGATGAGATC 3222

Qy      681 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3223 AAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAAGGTCTCCCTGAGAAATAC 3282

Qy      701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3283 AACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCT 3342

Qy      721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3343 ATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCC 3402

Qy      741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3403 CTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGATAAAGCCAGGACGGGAAGG 3462

Qy      761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3463 ACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGGTT 3522

Qy      781 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3523 CTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAGCTCCTGAGAAATCGAGAC 3582

Qy      801 IleTyrPheLysLeuValAsnAlaGlnSerValGln 812
            ||||||||||||||||||||||||||||||||||||
Db     3583 ATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3618
```

```
RESULT 4
US-09-873-409-13
; Sequence 13, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
;   LENGTH: 3702
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: No. US20020037522A1e
;   LOCATION: (723)..(723)
;   OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-13

Alignment Scores:
Pred. No.:              0              Length:         3702
```

```
Score:                     4079.00       Matches:        812
Percent Similarity:        100.00%       Conservative:   0
Best Local Similarity:     100.00%       Mismatches:     0
Query Match:               100.00%       Indels:         0
DB:                        10            Gaps:           0

US-09-873-409-2 (1-812) x US-09-873-409-13 (1-3702)

Qy        1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1264 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 1323

Qy       21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1324 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 1383

Qy       41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1384 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 1443

Qy       61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1444 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 1503

Qy       81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1504 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 1563

Qy      101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1564 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 1623

Qy      121 AlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThr 140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1624 GCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTAGCACACCGACTTTCTACT 1683

Qy      141 IleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAla 160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1684 ATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATGCTGGCGGAGAAAGGAGCA 1743

Qy      161 HisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1744 CATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTTGTGATGTCACAGGATATT 1803

Qy      181 LysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSer 200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1804 AAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCA 1863

Qy      201 LeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSer 220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1864 CTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCC 1923

Qy      221 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1924 ACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAAC 1983

Qy      241 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1984 AAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCTGTTCTAAATGGAACTGTT 2043

Qy      261 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2044 CATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATGTTTGGAAATAATGATAAA 2103

Qy      281 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2104 ACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTCGTCATTTTGGGTGTTATT 2163

Qy      301 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2164 TGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGAGCAGGGGAAATTTTAACG 2223

Qy      321 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2224 ATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAGGATATTGCCTGGTTTGAT 2283

Qy      341 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2284 GAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCCATAGATATAGCACAAATT 2343

Qy      361 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2344 CAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAATGCAACTAACATGGGACTT 2403

Qy      381 SerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAla 400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2404 TCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTCCTGATTCTGAGTATTGCT 2463

Qy      401 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2464 CCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATGACTGGATTTGCCAACAAA 2523

Qy      421 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2524 GATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAGAATATACGT 2583

Qy      441 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2584 ACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATGTATGAAGAGATGCTTCAG 2643

Qy      461 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2644 ACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTATGCATTCAGC 2703

Qy      481 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2704 CATGCCTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTTGGAGCCTATTTAATTCAA 2763

Qy      501 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2764 GCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACTGCAATTGCATATGGAGCT 2823

Qy      521 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2824 ATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCCAAAGCCAAATCGGGGGCT 2883

Qy      541 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2884 GCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGACAGCCGCAGTCAAGAAGGG 2943

Qy      561 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 580
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2944 AAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTCTTCTATCCA 3003

Qy      581 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3004 TGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACA 3063

Qy      601 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3064 GTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCTGTTCAACTTCTGCAGAGA 3123

Qy      621 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3124 CTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTGGATGCAAAAGAATTGAAT 3183

Qy      641 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3184 GTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAGCCTGTGCTCTTCAACTGC 3243

Qy      661 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3244 AGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTGGTGCCATTAGATGAGATC 3303

Qy      681 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     3304 AAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAAGGTCTCCCTGAGAAATAC 3363

Qy      701 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3364 AACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCT 3423

Qy      721 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3424 ATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCC 3483

Qy      741 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3484 CTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGATAAAGCCAGGACGGGAAGG 3543

Qy      761 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3544 ACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGGTT 3603

Qy      781 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3604 CTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAGCTCCTGAGAAATCGAGAC 3663

Qy      801 IleTyrPheLysLeuValAsnAlaGlnSerValGln 812
            ||||||||||||||||||||||||||||||||||||
Db     3664 ATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3699

RESULT 5
US-09-873-409-9
; Sequence 9, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
;   LENGTH: 2066
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-409-9

Alignment Scores:
Pred. No.:                0              Length:          2066
Score:                    3456.00        Matches:         687
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              84.73%         Indels:          0
DB:                       10             Gaps:            0

US-09-873-409-2 (1-812) x US-09-873-409-9 (1-2066)

Qy      126 SerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAsp 145
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        3 AGCAAAGGTCGGACTACAATCGTGGTAGCACACCGACTTTCTACTATTCGAAGTGCAGAT 62

Qy      146 LeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMet 165
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       63 TTGATTGTGACCCTAAAGGATGGAATGCTGGCGGAGAAAGGAGCACATGCTGAACTAATG 122

Qy      166 AlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGlu 185
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      123 GCAAAACGAGGTCTATATTATTCACTTGTGATGTCACAGGATATTAAAAAAGCTGATGAA 182

Qy      186 GlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSer 205
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      183 CAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCACTTCCTCTGCACTCT 242

Qy      206 ValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGlu 225
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      243 GTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCCACCCAATCTAAAGAG 302

Qy      226 IleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpPro 245
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      303 ATAAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAACAAGCCTGAATGGCCT 362

Qy      246 PheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSer 265
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      363 TTTGTGGTTCTGGGGACATTGGCTTCTGTTCTAAATGGAACTGTTCATCCAGTATTTTCC 422

Qy      266 IleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHis 285
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      423 ATCATCTTTGCAAAAATTATAACCATGTTTGGAAATAATGATAAAACCACATTAAAGCAT 482

Qy      286 AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr 305
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      483 GATGCAGAAATTTATTCCATGATATTCGTCATTTTGGGTGTTATTTGCTTTGTCAGTTAT 542

Qy      306 PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHis 325
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      543 TTCATGCAGGGATTATTTTACGGCAGAGCAGGGGAAATTTTAACGATGAGATTAAGACAC 602

Qy      326 LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSer 345
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      603 TTGGCCTTCAAAGCCATGTTATATCAGGATATTGCCTGGTTTGATGAAAAGGAAAACAGC 662

Qy      346 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      663 ACAGGAGGCTTGACAACAATATTAGCCATAGATATAGCACAAATTCAAGGAGCAACAGGT 722

Qy      366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      723 TCCAGGATTGGCGTCTTAACACAAAATGCAACTAACATGGGACTTTCAGTTATCATTTCC 782

Qy      386 PheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      783 TTTATATATGGATGGGAGATGACATTCCTGATTCTGAGTATTGCTCCAGTACTTGCCGTG 842

Qy      406 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      843 ACAGGAATGATTGAAACCGCAGCAATGACTGGATTTGCCAACAAAGATAAGCAAGAACTT 902

Qy      426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      903 AAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAGAATATACGTACTATAGTGTCATTA 962

Qy      446 ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn 465
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      963 ACAAGGGAAAAAGCCTTCGAGCAAATGTATGAAGAGATGCTTCAGACTCAACACAGAAAT 1022

Qy      466 ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr 485
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1023 ACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTATGCATTCAGCCATGCCTTTATATAT 1082

Qy      486 PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr 505
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1083 TTTGCCTATGCAGCAGGGTTTCGATTTGGAGCCTATTTAATTCAAGCTGGACGAATGACC 1142

Qy      506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys 525
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1143 CCAGAGGGCATGTTCATAGTTTTTACTGCAATTGCATATGGAGCTATGGCCATCGGAAAA 1202

Qy      526 ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAla 545
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1203 ACGCTCGTTTTGGCTCCTGAATATTCCAAAGCCAAATCGGGGGCTGCGCATCTGTTTGCC 1262

Qy      546 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 565
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1263 TTGTTGGAAAAGAAACCAAATATAGACAGCCGCAGTCAAGAAGGGAAAAAGCCAGACACA 1322

Qy      566 CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspVal 585
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     1323 TGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTCTTCTATCCATGTCGCCCAGATGTT 1382

Qy      586 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1383 TTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGCATTTGTGGGG 1442

Qy      606 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProVal 625
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1443 AGCAGCGGCTGTGGGAAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTG 1502

Qy      626 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArg 645
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1503 CAAGGACAAGTGCTGTTTGATGGTGTGGATGCAAAAGAATTGAATGTACAGTGGCTCCGT 1562

Qy      646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1563 TCCCAAATAGCAATCGTTCCTCAAGAGCCTGTGCTCTTCAACTGCAGCATTGCTGAGAAC 1622

Qy      666 IleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn 685
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1623 ATCGCCTATGGTGACAACAGCCGTGTGGTGCCATTAGATGAGATCAAAGAAGCCGCAAAT 1682

Qy      686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 705
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1683 GCAGCAAATATCCATTCTTTTATTGAAGGTCTCCCTGAGAAATACAACACACAAGTTGGA 1742

Qy      706 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1743 CTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTT 1802

Qy      726 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1803 CTCCAAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGATAATGACAGT 1862

Qy      746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1863 GAGAAGGTGGTTCAGCATGCCCTTGATAAAGCCAGGACGGGAAGGACATGCCTAGTGGTC 1922

Qy      766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1923 ACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGGTTCTGCACAATGGAAAG 1982

Qy      786 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 805
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1983 ATAAAGGAACAAGGAACTCATCAAGAGCTCCTGAGAAATCGAGACATATATTTTAAGTTA 2042

Qy      806 ValAsnAlaGlnSerValGln 812
            |||||||||||||||||||||
Db     2043 GTGAATGCACAGTCAGTGCAG 2063
```

RESULT 6
US-09-880-107-2299
; Sequence 2299, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2299
; LENGTH: 3924
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M23234
US-09-880-107-2299

Alignment Scores:			
Pred. No.:	5.86e-233	Length:	3924
Score:	2465.50	Matches:	478
Percent Similarity:	76.46%	Conservative:	149
Best Local Similarity:	58.29%	Mismatches:	182
Query Match:	60.44%	Indels:	11
DB:	10	Gaps:	3

US-09-873-409-2 (1-812) x US-09-880-107-2299 (1-3924)

```
Qy        2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
             :::|||      ||||||||     ||||||   :::   |||:::   ||||||||
Db     1395 ATTGATGGGCAGGATATTAGGAACTTTAATGTAAACTATCTGAGGGAAATCATTGGTGTG 1454

Qy       22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
            |||||||||||||||||||||||   |||||||||:::   ||||||   ||||||||||
Db     1455 GTGAGTCAGGAGCCGGTGCTGTTTTCCACCACAATTGCTGAAAATATTTGTTATGGCCGT 1514

Qy       42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
             :::||||||   :::|||:::::::::|||   :::||||||||||||||||:::|||
Db     1515 GGAAATGTAACCATGGATGAGATAAAGAAAGCTGTCAAAGAGGCCAACGCCTATGAGTTT 1574

Qy       62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
            ||||||:::   |||    ||||||:::||||||||||||:::|||||||||:::|||
Db     1575 ATCATGAAATTACCACAGAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGCTGAGT 1634

Qy       82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1635 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTT 1694

Qy      102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
            :::|||||||||||||||||||||||||||:::||||||::::::   ||||||||||||
Db     1695 CTGCTGGATGAGGCCACGTCAGCATTGGACACAGAAAGTGAAGCTGAGGTACAGGCAGCT 1754

Qy      122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
            |||:::||||||   :::||||||||||||||||||:::||||||||||||||||||:::
Db     1755 CTGGATAAGGCCAGAGAAGGCCGGACCACCATTGTGATAGCACACCGACTGTCTACGGTC 1814

Qy      142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
            |||:::||||||:::|||            :::||||||::::::   |||:::|||:::|||
Db     1815 CGAAATGCAGATGTCATCGCTGGGTTTGAGGATGGAGTAATTGTGGAGCAAGGAAGCCAC 1874

Qy      162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal------------------ 175
            :::|||||||||   |||   |||:::|||:::   ||||||
Db     1875 AGCGAACTGATGAAGAAGGAAGGGGTGTACTTCAAACTTGTCAACATGCAGACATCAGGA 1934

Qy      176 ---------MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyr 192
                     |||::::::   :::      ||||||:::         :::
Db     1935 AGCCAGATCCAGTCAGAAGAATTTGAACTAAATGATGAAAAGGCTGCCACTAGAATGGCC 1994

Qy      193 SerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys---Ser 211
                     |||:::                        :::   |||::::::|||   |||
Db     1995 CCAAATGGCTGGAAATCTCGCCTATTTAGGCATTCTACTCAGAAAAACCTTAAAAATTCA 2054

Qy      212 AspPheIleAspLysAla---GluGluSerThrGlnSerLysGluIleSerLeuProGlu 230
                     |||:::   :::      |||      |||      ::::::|||
Db     2055 CAAATGTGTCAGAAGAGCCTTGATGTGGAAACCGATGGACTTGAAGCAAATGTGCCACCA 2114

Qy      231 ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGly 250
            ||||||   ||||||:::||||||||||||||||||   |||||||||:::   |||:::|||
Db     2115 GTGTCCTTTCTGAAGGTCCTGAAACTGAATAAAACAGAATGGCCCTACTTTGTCGTGGGA 2174

Qy      251 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 270
            |||:::   ::::::   ||||||   :::   |||   ||||||:::||||||::::::
Db     2175 ACAGTATGTGCCATTGCCAATGGGGGGCTTCAGCCGGCATTTTCAGTCATATTCTCAGAG 2234

Qy      271 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 290
            ||||||   :::||||||         |||                  :::         |||:::
Db     2235 ATCATAGCGATTTTTGGACCAGGCGATGATGCAGTGAAGCAGCAGAAGTGCAACATATTC 2294
```

```
Qy    291 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 310
Db   2295 TCTTTGATTTTCTTATTTCTGGGAATTATTTCTTTTTTACTTTCTTCCTTCAGGGTTTC 2354

Qy    311 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 330
Db   2355 ACGTTTGGGAAAGCTGGCGAGATCCTCACCAGAAGACTGCGGTCAATGGCTTTTAAAGCA 2414

Qy    331 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 350
Db   2415 ATGCTAAGACAGGACATGAGCTGGTTTGATGACCATAAAAACAGTACTGGTGCACTTTCT 2474

Qy    351 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 370
Db   2475 ACAAGACTTGCCACAGATGCTGCCCAAGTCCAAGGAGCCACAGGAACCAGGTTGGCTTTA 2534

Qy    371 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrp 390
Db   2535 ATTGCACAGAATATAGCTAACCTTGGAACTGGTATTATCATATCATTTATCTACGGTTGG 2594

Qy    391 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 410
Db   2595 CAGTTAACCCTATTGCTATTAGCAGTTGTTCCAATTATTGCTGTGTCAGGAATTGTTGAA 2654

Qy    411 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 430
Db   2655 ATGAAATTGTTGGCTGGAAATGCCAAAAGAGATAAAAAAGAACTGGAAGCTGCTGGAAAG 2714

Qy    431 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 450
Db   2715 ATTGCAACAGAGGCAATAGAAAATATTAGGACAGTTGTGTCTTTGACCCAGGAAAGAAAA 2774

Qy    451 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 470
Db   2775 TTTGAATCAATGTATGTTGAAAAATTGTATGGACCTTACAGGAATTCTGTGCAGAAGGCA 2834

Qy    471 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 490
Db   2835 CACATCTATGGAATTACTTTTAGTATCTCACAAGCATTTATGTATTTTTCCTATGCCGGT 2894

Qy    491 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 510
Db   2895 TGTTTTCGATTTGGTGCATATCTCATTGTGAATGGACATATGCGCTTCAGAGATGTTATT 2954

Qy    511 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 530
Db   2955 CTGGTGTTTTCTGCAATTGTATTTGGTGCAGTGGCTCTAGGACATGCCAGTTCATTTGCT 3014

Qy    531 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 550
Db   3015 CCAGACTATGCTAAAGCTAAGCTGTCTGCAGCCCACTTATTCATGCTGTTTGAAAGACAA 3074

Qy    551 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 570
Db   3075 CCTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAGCCTGATAAATTTGAAGGAAATATA 3134

Qy    571 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 590
Db   3135 ACATTTAATGAAGTCGTGTTCAACTATCCCACCCGAGCAAACGTGCCAGTGCTTCAGGGG 3194

Qy    591 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 610
Db   3195 CTGAGCCTGGAGGTGAAGAAAGGCCAGACACTAGCCCTGGTGGGCAGCAGTGGCTGTGGG 3254

Qy    611 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 630
Db   3255 AAGAGCACGGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCGGGGACAGTGCTT 3314

Qy    631 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 650
Db   3315 CTCGATGGTCAAGAAGCAAAGAAACTCAATGTCCAGTGGCTCAGAGCTCAACTCGGAATC 3374

Qy    651 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 670
Db   3375 GTGTCTCAGGAGCCTATCCTATTTGACTGCAGCATTGCCGAGAATATTGCCTATGGAGAC 3434

Qy    671 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 690
Db   3435 AACAGCCGGGTTGTATCACAGGATGAAATTGTGAGTGCAGCCAAAGCTGCCAACATACAT 3494

Qy    691 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 710
Db   3495 CCTTTCATCGAGACGTTACCCCACAAATATGAAACAAGAGTGGGAGATAAGGGGACTCAG 3554

Qy    711 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 730
Db   3555 CTCTCAGGAGGTCAAAAACAGAGGATTGCTATTGCCCGAGCCCTCATCAGACAACCTCAA 3614

Qy    731 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 750
Db   3615 ATCCTCCTGTTGGATGAAGCTACATCAGCTCTGGATACTGAAAGTGAAAAGGTTGTCCAA 3674

Qy    751 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 770
Db   3675 GAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCC 3734

Qy    771 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 790
Db   3735 ACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGGAGAGTCAAGGAGCATGGC 3794

Qy    791 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 810
Db   3795 ACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3854
```

```
RESULT 7
US-09-769-097-1
; Sequence 1, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
;  APPLICANT: Kimberly Anne Brun
;  APPLICANT:  Richard James Chenery
;  APPLICANT:  Harma Ellens
;  APPLICANT:  John Anthony Feild
;  APPLICANT:  Lin Yue
;  TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
;  TITLE OF INVENTION:  ENCODING RAT MDR1A AND SCREENING METHODS THEREOF
;  FILE REFERENCE: GP-50009-C2
;  CURRENT APPLICATION NUMBER: US/09/769,097
;  CURRENT FILING DATE:  2001-01-25
;  PRIOR APPLICATION NUMBER: 09/208,809
;  PRIOR FILING DATE: 1998-12-09
;  PRIOR APPLICATION NUMBER: 09/156,800
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: US99/20770
;  PRIOR FILING DATE: 1999-09-10
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 4369
;   TYPE: DNA
;   ORGANISM: RATTUS RATTUS
US-09-769-097-1

Alignment Scores:
Pred. No.:               1.79e-230      Length:         4369
Score:                   2441.00        Matches:        470
Percent Similarity:      77.04%         Conservative:   164
Best Local Similarity:   57.11%         Mismatches:     175
Query Match:             59.84%         Indels:         14
DB:                      10             Gaps:           5

US-09-873-409-2 (1-812) x US-09-769-097-1 (1-4369)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
```

```
Db   1684 ATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTG 1743

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db   1744 GTGAGTCAGGAACCCGTGCTGTTTGCCACCACAATTGCCGAAAACATTCGCTATGGCCGA 1803

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db   1804 GAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAATGCCTATGATTTC 1863

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db   1864 ATCATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGT 1923

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db   1924 GGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAACCCCAAGATCCTT 1983

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db   1984 TTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTGGTTCAGGCCGCT 2043

Qy    122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db   2044 CTGGATAAGGCTAGAGAAGGCCGGACCACCATTGTGATAGCTCACCGCTTGTCTACAGTT 2103

Qy    142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db   2104 CGCAATGCTGACGTCATTGCTGGTTTTGATGGTGGTGTCATTGTGGAGCAAGGAAATCAT 2163

Qy    162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178
Db   2164 GATGAGCTCATGAGAGAGAAAGGAATTTACTTCAAACTTGTCATGACTCAGACAGCAGGA 2223

Qy    179 ---AspIleLysLysAlaAspGluGlnMetGluSer-----------------MetThr 191
Db   2224 AATGAAATTGAATTAGGAAATGAAGCTTGTGAATCTAAAGACGGAATTGATAATGTGGAC 2283

Qy    192 TyrSerThrGluArgLysThrAsnSerLeu------ProLeuHisSerValLysSerIle 209
Db   2284 ATGTCTTCAAAAGATTCGGGATCCAGTCTAATAAGAAGAAGATCAACTCGCAAAAGCATC 2343

Qy    210 LysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLysGluIleSerLeu 228
Db   2344 CGTGGGCCACATGATCAAGACGGGGAACTTAGCACCAAAGAGGCTCTGGATGACGACGTA 2403

Qy    229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
Db   2404 CCTCCAGCTTCCTTTTGGCGGATCCTGAAGTTGAATTCAACTGAATGGCCTTATTTTGTG 2463

Qy    249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db   2464 GTTGGTGTATTTTGTGCCATAATAAATGGAGGCTTGCAACCAGCATTCTCCATAATATTT 2523

Qy    269 AlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeu---LysHisAspAla 287
Db   2524 TCAAAGGTTGTAGGGGTTTTACAAAAAATGACACCCCTGAAATCCAGCGGCAGAACAGC 2583

Qy    288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db   2584 AACTTGTTTTCTTTATTGTTCCTGATCCTTGGGATCATCTCTTTCATTACGTTTTTCCTT 2643

Qy    308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db   2644 CAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGACTCCGATACATGGTC 2703

Qy    328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
Db   2704 TTCAAATCCATGCTGAGACAGGACATAAGCTGGTTTGATGACCCTAAAAACACCACAGGA 2763

Qy    348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db   2764 GCGCTGACCACCAGGCTTGCCAATGACGCTGCTCAAGTGAAAGGGGCTACAGGGTCTAGG 2823

Qy    368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db   2824 CTTGCTGTTATTACCCAGAACATAGCAAATCTTGGGACAGGCATCATCATATCCCTGATC 2883

Qy    388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db   2884 TACGGCTGGCAATTGACACTTTTACTCCTAGCAATTGTTCCCATCATTGCTATAGCAGGA 2943

Qy    408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db   2944 GTGGTTGAAATGAAAATGTTGTCTGGACAAGCGCTGAAAGATAAGAAGGAACTAGAAGGT 3003

Qy    428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db   3004 TCTGGGAAGATCGCTACAGAAGCAATTGAAAACTTTCGCACTGTCGTCTCTTTGACTCGG 3063

Qy    448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db   3064 GAGCAGAAGTTTGAAACTATGTATGCCCAGAGCTTGCAGATACCATACAGAAATGCTTTG 3123

Qy    468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db   3124 AAGAAAGCGCACGTCTTTGGGATCACTTTCTCCTTCACCCAGGCCATGATGTATTTCTCC 3183

Qy    488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db   3184 TATGCTGCTTGTTTCCGGTTTGATGCCTACTTGGTGGCACGAGAACTCATGACATTTGAA 3243

Qy    508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db   3244 AATGTTCTGTTAGTATTCTCAGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAGGTCAGT 3303

Qy    528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db   3304 TCATTCGCTCCTGACTACGCGAAAGCCAAAGTCTCGGCATCCCACATCATCAGGATCATT 3363

Qy    548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db   3364 GAGAAAATCCCTGAGATTGACAGCTACAGCACGGAGGGCTTGAAGCCTAATATGTTGGAA 3423

Qy    568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db   3424 GGAAATGTGAAATTTAATGGAGTCATGTTCAACTATCCCACCCGACCCAACATCCCAGTG 3483

Qy    588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db   3484 CTTCAGGGGCTGAGCCTAGAGGTGAAGAAAGGGCAGACGCTGGCCCTCGTGGGCAGCAGT 3543

Qy    608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db   3544 GGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTTGAGCGCTTCTATGACCCCATGGCCGGA 3603

Qy    628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
Db   3604 ACAGTGTTTCTAGATGGCAAAGAAATAAAGCAACTCAATGTCCAGTGGCTCCGCGCCCAC 3663

Qy    648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db   3664 CTGGGCATTGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATCGCCGAGAACATTGCC 3723

Qy    668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db   3724 TACGGAGACAACAGCCGTGTCGTGTCTCATAAGGAGATCGTGAAGGCAGCCAAGGAGGCC 3783

Qy    688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db   3784 AACATCCACCAGTTCATCGACTCACTGCCTGAGAAATACAACACCAGAGTGGGAGACAAA 3843

Qy    708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
Db   3844 GGGACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATCGCCATCGCGCGCGCCCTCGTCAGA 3903
```

```
Qy     728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db    3904 CAGCCTCACATCTTACTTCTGGATGAAGCGACATCAGCTCTGGATACGGAGAGTGAAAAG 3963

Qy     748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db    3964 GTCGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACCTGCATTGTGATCGCGCAC 4023

Qy     768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db    4024 CGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATTCAGAACGGCCAGGTCAAG 4083

Qy     788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db    4084 GAGCACGGCACCCACCAGCAGCTGCTGGCCCAGAAAGGCATCTATTTCTCGATGGTCAGT 4143

Qy     808 AlaGlnSer 810
Db    4144 GTGCAGGCT 4152
```

RESULT 8
US-09-769-097-3
; Sequence 3, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: ENCODING RAT MDR1A AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP-50009-C2
; CURRENT APPLICATION NUMBER: US/09/769,097
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 4425
; TYPE: DNA
; ORGANISM: RATTUS RATTUS
US-09-769-097-3

Alignment Scores:

Pred. No.:	1.82e-230	Length:	4425
Score:	2441.00	Matches:	470
Percent Similarity:	77.04%	Conservative:	164
Best Local Similarity:	57.11%	Mismatches:	175
Query Match:	59.84%	Indels:	14
DB:	10	Gaps:	5

US-09-873-409-2 (1-812) x US-09-769-097-3 (1-4425)

```
Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db    1684 ATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTG 1743

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db    1744 GTGAGTCAGGAACCCGTGCTGTTTGCCACCACAATTGCCGAAAACATTCGCTATGGCCGA 1803

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db    1804 GAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAATGCCTATGATTTC 1863

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db    1864 ATCATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGT 1923

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db    1924 GGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAACCCCAAGATCCTT 1983

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db    1984 TTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTGGTTCAGGCCGCT 2043

Qy     122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db    2044 CTGGATAAGGCTAGAGAAGGCCGGACCACCATTGTGATAGCTCACCGCTTGTCTACAGTT 2103

Qy     142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db    2104 CGCAATGCTGAYGTCATTGCTGGTTTTGATGGTGGTGTCATTGTGGAGCAAGGAAATCAT 2163

Qy     162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178
Db    2164 GATGAGCTCATGAGAGAGAAAGGAATTTACTTCAAACTTGTCATGACTCAGACAGCAGGA 2223

Qy     179 ---AspIleLysLysAlaAspGluGlnMetGluSer-----------------MetThr 191
Db    2224 AATGAAATTGAATTAGGAAATGAAGCTTGTGAATCTAAAGAYGGAATTGATAATGTGGAC 2283

Qy     192 TyrSerThrGluArgLysThrAsnSerLeu------ProLeuHisSerValLysSerIle 209
Db    2284 ATGTCTTCAAAAGATTCRGGATCCAGTCTAATAAGAAGAAGATCAACTCGCAAAAGCATC 2343

Qy     210 LysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLysGluIleSerLeu 228
Db    2344 CGTGGGCCACATGATCAAGACGGGGAACTTAGCACCAAAGAGGCTCTGGATGACGACGTA 2403

Qy     229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
Db    2404 CCTCCAGCTTCCTTTTGGCGGATCCTGAAGTTGAATTCAACTGAATGGCCTTATTTTGTG 2463

Qy     249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db    2464 GTTGGTGTATTTTGTGCCATAATAAATGGAGGCTTGCAACCAGCATTCTCCATAATATTT 2523

Qy     269 AlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeu---LysHisAspAla 287
Db    2524 TCAAAGGTTGTAGGGGTTTTACAAAAAATGACACCCCTGAAATCCAGCGGCAGAACAGC 2583

Qy     288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db    2584 AACTTGTTTTCTTTATTGTTCCTGATCCTTGGGATCATCTCTTTCATTACGTTTTTCCTT 2643

Qy     308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db    2644 CAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGACTCCGATACATGGTC 2703

Qy     328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
Db    2704 TTCAAATCCATGCTGAGACAGGACATAAGCTGGTTTGATGACCCTAAAAACACCACAGGA 2763

Qy     348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db    2764 GCGCTGACCACCAGGCTTGCCAATGACGCTGCTCAAGTGAAAGGGGCTACAGGGTCTAGG 2823

Qy     368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db    2824 CTTGCTGTTATTACCCAGAACATAGCAAATCTTGGGACAGGCATCATCATATCCCTGATC 2883

Qy     388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db    2884 TACGGCTGGCAATTGACACTTTTACTCCTAGCAATTGTTCCCATCATTGCTATAGCAGGA 2943

Qy     408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db    2944 GTGGTTGAAATGAAAATGTTGTCTGGACAAGCGCTGAAAGATAAGAAGGAACTAGAAGGT 3003
```

```
Qy    428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db   3004 TCTGGGAAGATCGCTACAGAAGCAATTGAAAACTTTCGCACTGTCGTCTCTTTGACTCGG 3063

Qy    448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db   3064 GAGCAGAAGTTTGAAACTATGTATGCCCAGAGCTTGCAGATACCATACAGAAATGCTTTG 3123

Qy    468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db   3124 AAGAAAGCGCACGTCTTTGGGATCACTTTCTCCTTCACCCAGGCCATGATGTATTTCTCC 3183

Qy    488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db   3184 TATGCTGCTTGTTTCCGGTTTGATGCCTACTTGGTGGCACGAGAACTCATGACATTTGAA 3243

Qy    508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db   3244 AATGTTCTGTTAGTATTCTCAGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAGGTCAGT 3303

Qy    528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db   3304 TCATTCGCTCCTGACTACGCGAAAGCCAAAGTCTCGGCATCCCACATCATCAGGATCATT 3363

Qy    548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db   3364 GAGAAAATCCCTGAGATTGACAGCTACAGCACGGAGGGCTTGAAGCCTAATATGTTGGAA 3423

Qy    568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db   3424 GGAAATGTGAAATTTAATGGAGTCATGTTCAACTATCCCACCCGACCCAACATCCCAGTG 3483

Qy    588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db   3484 CTTCAGGGGCTGAGCCTAGAGGTGAAGAAAGGGCAGACGCTGGCCCTCGTGGGCAGCAGT 3543

Qy    608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db   3544 GGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTTGAGCGCTTCTATGACCCCATGGCCGGA 3603

Qy    628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
Db   3604 ACAGTGTTTCTAGATGGCAAAGAAATAAAGCAACTCAATGTCCAGTGGCTCCGCGCCCAC 3663

Qy    648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db   3664 CTGGGCATTGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATCGCCGAGAACATTGCC 3723

Qy    668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db   3724 TACGGAGACAACAGCCGTGTCGTGTCTCATAAGGAGATCGTGAAGGCAGCCAAGGAGGCC 3783

Qy    688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db   3784 AACATCCACCAGTTCATCGACTCACTGCCTGAGAAATACAACACCAGAGTGGGAGACAAA 3843

Qy    708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
Db   3844 GGGACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATCGCCATCGCGCGCGCCCTCGTCAGA 3903

Qy    728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db   3904 CAGCCTCACATCTTACTTCTGGATGAAGCGACATCAGCTCTGGATACGGAGAGTGAAAAG 3963

Qy    748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db   3964 GTCGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACCTGCATTGTGATCGCGCAC 4023

Qy    768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db   4024 CGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATTCAGAACGGCCAGGTCAAG 4083

Qy    788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db   4084 GAGCACGGCACCCACCAGCAGCTGCTGGCCCAGAAAGGCATCTATTTCTCGATGGTCAGT 4143

Qy    808 AlaGlnSer 810
Db   4144 GTGCAGGCT 4152
```

RESULT 9
US-10-044-671-1
; Sequence 1, Application US/10044671
; Patent No. US20020177147A1
; GENERAL INFORMATION:
; APPLICANT: Washington State University Research Foundation
; APPLICANT: Mealey, Katrina
; APPLICANT: Bentjen, Steven
; TITLE OF INVENTION: MDR1 VARIANTS AND METHODS FOR THEIR USE
; FILE REFERENCE: 4630-61733
; CURRENT APPLICATION NUMBER: US/10/044,671
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/261,578
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/314,829
; PRIOR FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4317
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (70)..(3912)
; OTHER INFORMATION:
US-10-044-671-1

Alignment Scores:

Pred. No.:	3.47e-230	Length:	4317
Score:	2438.00	Matches:	467
Percent Similarity:	76.24%	Conservative:	162
Best Local Similarity:	56.61%	Mismatches:	178
Query Match:	59.77%	Indels:	18
DB:	9	Gaps:	4

US-09-873-409-2 (1-812) x US-10-044-671-1 (1-4317)

```
Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db   1429 ATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTG 1488

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db   1489 GTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGC 1548

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db   1549 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTT 1608

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db   1609 ATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAGAGAGGGGCCCGGCTGAGT 1668

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db   1669 GGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTT 1728

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db   1729 CTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCC 1788

Qy    122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db   1789 CTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCTCATCGTTTGTCTACAGTT 1848
```

```
Qy    142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db   1849 CGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 1908

Qy    162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLys 181
Db   1909 GATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTCACAATGCAG---ACAAGA 1965

Qy    182 LysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu 201
Db   1966 GGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAAAGTGAAAGTGATGCCTTG 2025

Qy    202 ProLeuHisSerValLysSerIleLysSerAspPheIle--------------------- 214
Db   2026 GAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAAAGAAGATCAACTCGCAGG 2082

Qy    215 ------------------------AspLysAlaGluGluSerThrGlnSerLysGluIle 226
Db   2083 AGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACAAAAGAGGACTTGAATGAG 2142

Qy    227 SerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPhe 246
Db   2143 AATGTTCCTTCAGTTTCCTTCTGGAGGATTCTGAAGCTGAACTCAACTGAATGGCCTTAT 2202

Qy    247 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 266
Db   2203 TTTGTGGTTGGTATATTTTGTGCTATTATAAACGGAGGCCTGCAACCAGCATTTTCAATA 2262

Qy    267 IlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHis 285
Db   2263 ATATTTTCAAGGATTATAGGGATCTTTACCCGAGATGAGGATCCTGAAACAAAACGACAG 2322

Qy    286 AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr 305
Db   2323 AATAGTAACATGTTTTCTGTATTGTTTCTAGTCCTTGGAATTATTTCTTTTATTACATTT 2382

Qy    306 PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHis 325
Db   2383 TTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTCACTAAGCGGCTTCGATAC 2442

Qy    326 LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSer 345
Db   2443 ATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTTGATGACCCTAAAAACACC 2502

Qy    346 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 365
Db   2503 ACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAAGTTAAAGGGGCTATAGGT 2562

Qy    366 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 385
Db   2563 TCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGGACAGGCATTATTATATCC 2622

Qy    386 PheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 405
Db   2623 TTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATTGTACCCATCATTGCAATA 2682

Qy    406 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 425
Db   2683 GCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAGCTA 2742

Qy    426 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 445
Db   2743 GAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTCCGAACTGTTGTTTCTTTG 2802

Qy    446 ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn 465
Db   2803 ACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTGCAAGTACCATACAGAAAC 2862

Qy    466 ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr 485
Db   2863 TCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATCACCCAGGCAATGATGTAT 2922

Qy    486 PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr 505
Db   2923 TTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTGGCAAATGAGTTCATGAAC 2982

Qy    506 ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys 525
Db   2983 TTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAG 3042

Qy    526 ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAla 545
Db   3043 GTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCAGCAGCCCACGTCATCATG 3102

Qy    546 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 565
Db   3103 ATCATTGAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCACGGCCTCAAGCCAAATACG 3162

Qy    566 CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspVal 585
Db   3163 TTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTATCCCACTCGACCAGACATC 3222

Qy    586 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 605
Db   3223 CCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAGACGCTGGCCCTCGTAGGT 3282

Qy    606 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProVal 625
Db   3283 AGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAGCGCTTCTATGACCCCTTG 3342

Qy    626 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArg 645
Db   3343 GCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTGAATGTCCAGTGGCTCCGA 3402

Qy    646 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 665
Db   3403 GCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGACTGCAGCATTGCCGAGAAC 3462

Qy    666 IleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn 685
Db   3463 ATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAGATTATGCAGGCAGCCAAG 3522

Qy    686 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 705
Db   3523 GAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAATACAACACCAGAGTAGGA 3582

Qy    706 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 725
Db   3583 GACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATTGCCATAGCTCGCGCTCTT 3642

Qy    726 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 745
Db   3643 GTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAGT 3702

Qy    746 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 765
Db   3703 GAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATC 3762

Qy    766 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 785
Db   3763 GCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTGGTGTTTCAGAATGGCAAA 3822

Qy    786 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 805
Db   3823 GTCAAGGAGCATGGCACACATCAACAGCTGCTGGCTCAGAAAGGCATCTATTTTTCCATG 3882

Qy    806 ValAsnAlaGlnSer 810
Db   3883 ATCAGTGTCCAGGCT 3897
```

RESULT 10
US-09-866-866A-1
; Sequence 1, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:

```
;  APPLICANT: Sorrentino, Brian
;  APPLICANT:  Schuetz, John
;  TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
;  FILE REFERENCE: 1340-1-021CIP2
;  CURRENT APPLICATION NUMBER: US/09/866,866A
;  CURRENT FILING DATE:  2001-08-30
;  PRIOR APPLICATION NUMBER: 09/584,586
;  PRIOR FILING DATE: 2000-05-31
;  PRIOR APPLICATION NUMBER: PCT/US99/11825
;  PRIOR FILING DATE: 1999-05-27
;  PRIOR APPLICATION NUMBER: 60/086,988
;  PRIOR FILING DATE: 1998-05-28
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 3860
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-866A-1

Alignment Scores:
Pred. No.:               4.62e-230     Length:          3860
Score:                   2436.00       Matches:         466
Percent Similarity:      76.79%        Conservative:    166
Best Local Similarity:   56.62%        Mismatches:      177
Query Match:             59.72%        Indels:          14
DB:                      10            Gaps:            4

US-09-873-409-2 (1-812) x US-09-866-866A-1 (1-3860)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db    1357 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 1416

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db    1417 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 1476

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db    1477 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1536

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db    1537 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 1596

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db    1597 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 1656

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db    1657 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 1716

Qy     122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db    1717 CTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGTTTGTCTACAGTT 1776

Qy     142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db    1777 CGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 1836

Qy     162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178
Db    1837 GATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATGCAGACAGCAGGA 1896

Qy     179 ---------AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
Db    1897 AATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATTGATGCCTTGGAA 1956

Qy     196 ArgLysThrAsnSerLeuProLeuHisSerVal------------------LysSerIle 209
Db    1957 ATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGTC 2016

Qy     210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228
Db    2017 CGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTGGATGAAAGTATA 2076

Qy     229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
Db    2077 CCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCCTTATTTTGTT 2136

Qy     249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db    2137 GTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTTGCAATAATATTT 2196

Qy     269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
Db    2197 TCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAACGACAGAATAGT 2256

Qy     288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db    2257 AACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATTACATTTTTCCTT 2316

Qy     308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db    2317 CAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTT 2376

Qy     328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
Db    2377 TTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAAAACACCACTGGA 2436

Qy     348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db    2437 GCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGG 2496

Qy     368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db    2497 CTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATC 2556

Qy     388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db    2557 TATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGA 2616

Qy     408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db    2617 GTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGT 2676

Qy     428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db    2677 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAG 2736

Qy     448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db    2737 GAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 2796

Qy     468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db    2797 AGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCC 2856

Qy     488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db    2857 TATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAG 2916

Qy     508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db    2917 GATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGT 2976

Qy     528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db    2977 TCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATT 3036

Qy     548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db    3037 GAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAA 3096
```

```
Qy     568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
           |||||::::   |||   ||||||   |||   ||||||   ||||||||:::   :::
Db    3097 GGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTG 3156

Qy     588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
           |||:::||||||||||||   :::::::::|||:::|||:::|||   ||||||||||||
Db    3157 CTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGT 3216

Qy     608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
           ||||||||||||||||||   ||||||||||||:::|||   |||||||||:::   |||
Db    3217 GGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGG 3276

Qy     628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
           :::||||||   ||||||   :::   |||   ||||||||||||||||||||||||:::
Db    3277 AAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACAC 3336

Qy     648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
           :::   ||||||   |||||||||:::||||||:::||||||||||||||||||||||||
Db    3337 CTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCC 3396

Qy     668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
           ||||||||||||||||||||||||      :::||||||      ||||||      |||
Db    3397 TATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCC 3456

Qy     688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
           |||||||||:::|||||||||   ||||||   ||||||:::|||:::||||||   |||
Db    3457 AACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAA 3516

Qy     708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
           |||   |||||||||||||||||||||||||||:::||||||||||||||||||::::::
Db    3517 GGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGA 3576

Qy     728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
           :::|||   ||||||||||||||||||||||||||||||||||||   :::|||||||||
Db    3577 CAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAG 3636

Qy     748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
           |||||||||   ||||||||||||||||||||||||||||||:::|||:::   |||
Db    3637 GTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCAC 3696

Qy     768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
           |||||||||   |||||||||||||||||||||||||||   ||||||::::::|||
Db    3697 CGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGCAGAGTCAAG 3756

Qy     788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
           |||   ||||||||||||:::||||||   :::   |||||||||   :::|||:::
Db    3757 GAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGT 3816

Qy     808 AlaGlnSer 810
              |||:::
Db    3817 GTCCAGGCT 3825
```

RESULT 11
US-09-866-866A-3
; Sequence 3, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
;   LENGTH: 3860
;   TYPE: DNA
;   ORGANISM: homo sapiens
US-09-866-866A-3

Alignment Scores:

Pred. No.:	4.62e-230	Length:	3860
Score:	2436.00	Matches:	466
Percent Similarity:	76.79%	Conservative:	166
Best Local Similarity:	56.62%	Mismatches:	177
Query Match:	59.72%	Indels:	14
DB:	10	Gaps:	4

US-09-873-409-2 (1-812) x US-09-866-866A-3 (1-3860)

```
Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
           ||||||   |||||||||   :::|||||||||   ||||:::   ||||||||||
Db    1357 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 1416

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           ||||||||||||||||||||||||   |||||||||:::   ||||||:::|||||||||
Db    1417 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 1476

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
           ::::::||||||   :::|||:::|||:::|||   :::||||||||||||||||||||||
Db    1477 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1536

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           ||||||:::   |||:::||||||:::||||||||||||||||:::|||||||||:::|||
Db    1537 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 1596

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1597 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 1656

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
           :::||||||||||||||||||||||||||||:::||||||::::::   ||||||   |||
Db    1657 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 1716

Qy     122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
           |||:::||||||   ||||||||||||||||||||||:::|||||||||||||||||:::
Db    1717 CTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGTTTGTCTACAGTT 1776

Qy     142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
           |||:::||||||:::|||      ||||||::::::   |||||||||      |||
Db    1777 CGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 1836

Qy     162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178
              |||||||||   ::::::|||:::|||:::   ||||||      |||
Db    1837 GATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATGCAGACAGCAGGA 1896

Qy     179 ---------AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
                    :::::::::   |||   ::::::   :::|||         :::   |||
Db    1897 AATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATTGATGCCTTGGAA 1956

Qy     196 ArgLysThrAsnSerLeuProLeuHisSerVal-----------------LysSerIle 209
              :::|||         :::                         :::|||:::
Db    1957 ATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGTC 2016

Qy     210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228
           :::   |||      |||:::      :::   ::::::   :::   |||:::
Db    2017 CGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTGGATGAAAGTATA 2076

Qy     229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
           |||   ||||||      :::|||:::||||||      |||||||||:::   |||
Db    2077 CCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCCTTATTTTGTT 2136

Qy     249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
           :::|||      :::::::::||||||   :::   |||   |||:::|||||||||
Db    2137 GTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTTGCAATAATATTT 2196

Qy     269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
           :::|||||||||   :::|||         :::|||      |||   :::   ::::::
Db    2197 TCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAACGACAGAATAGT 2256
```

```
Qy    288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db   2257 AACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATTACATTTTTCCTT 2316

Qy    308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db   2317 CAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTT 2376

Qy    328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
Db   2377 TTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAAAACACCACTGGA 2436

Qy    348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db   2437 GCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGG 2496

Qy    368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db   2497 CTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATC 2556

Qy    388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db   2557 TATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGA 2616

Qy    408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db   2617 GTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGT 2676

Qy    428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db   2677 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAG 2736

Qy    448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db   2737 GAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 2796

Qy    468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db   2797 AGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCC 2856

Qy    488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db   2857 TATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAG 2916

Qy    508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db   2917 GATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGT 2976

Qy    528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db   2977 TCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATT 3036

Qy    548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db   3037 GAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAA 3096

Qy    568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db   3097 GGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTG 3156

Qy    588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db   3157 CTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGT 3216

Qy    608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db   3217 GGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGG 3276

Qy    628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
Db   3277 AAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACAC 3336

Qy    648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db   3337 CTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCC 3396

Qy    668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db   3397 TATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCC 3456

Qy    688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db   3457 AACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAA 3516

Qy    708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
Db   3517 GGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGA 3576

Qy    728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db   3577 CAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAG 3636

Qy    748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db   3637 GTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCAC 3696

Qy    768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db   3697 CGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGCAGAGTCAAG 3756

Qy    788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db   3757 GAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGT 3816

Qy    808 AlaGlnSer 810
Db   3817 GTCCAGGCT 3825
```

RESULT 12
US-10-072-621-2
; Sequence 2, Application US/10072621
; Patent No. US20020169137A1
; GENERAL INFORMATION:
; APPLICANT: Reiner, Peter B.
; APPLICANT: Connop, Bruce P.
; APPLICANT: Pollard, Michelle
; TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION
; TITLE OF INVENTION: BY MODIFICATION OF ABC TRANSPORTER EXPRESSION OR ACTIVITY
; FILE REFERENCE: 100103.402
; CURRENT APPLICATION NUMBER: US/10/072,621
; CURRENT FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 4643
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-072-621-2

Alignment Scores:

Pred. No.:	6.08e-230	Length:	4643
Score:	2436.00	Matches:	466
Percent Similarity:	76.79%	Conservative:	166
Best Local Similarity:	56.62%	Mismatches:	177
Query Match:	59.72%	Indels:	14
DB:	9	Gaps:	4

US-09-873-409-2 (1-812) x US-10-072-621-2 (1-4643)

```
Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db   1778 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 1837

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
```

```
Db   1838 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 1897

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db   1898 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1957

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db   1958 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 2017

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db   2018 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 2077

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db   2078 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 2137

Qy    122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db   2138 CTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGTTTGTCTACAGTT 2197

Qy    142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db   2198 CGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 2257

Qy    162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178
Db   2258 GATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATGCAGACAGCAGGA 2317

Qy    179 ---------AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
Db   2318 AATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATTGATGCCTTGGAA 2377

Qy    196 ArgLysThrAsnSerLeuProLeuHisSerVal------------------LysSerIle 209
Db   2378 ATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGTC 2437

Qy    210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228
Db   2438 CGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTGGATGAAAGTATA 2497

Qy    229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
Db   2498 CCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCCTTATTTTGTT 2557

Qy    249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db   2558 GTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTTGCAATAATATTT 2617

Qy    269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
Db   2618 TCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAACGACAGAATAGT 2677

Qy    288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db   2678 AACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATTACATTTTTCCTT 2737

Qy    308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db   2738 CAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTT 2797

Qy    328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
Db   2798 TTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAAAACACCACTGGA 2857

Qy    348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db   2858 GCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGG 2917

Qy    368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db   2918 CTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATC 2977

Qy    388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db   2978 TATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGA 3037

Qy    408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db   3038 GTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGT 3097

Qy    428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db   3098 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAG 3157

Qy    448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db   3158 GAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 3217

Qy    468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db   3218 AGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCC 3277

Qy    488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db   3278 TATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAG 3337

Qy    508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db   3338 GATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGT 3397

Qy    528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db   3398 TCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATT 3457

Qy    548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db   3458 GAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAA 3517

Qy    568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db   3518 GGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTG 3577

Qy    588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db   3578 CTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGT 3637

Qy    608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db   3638 GGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGG 3697

Qy    628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
Db   3698 AAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACAC 3757

Qy    648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db   3758 CTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCC 3817

Qy    668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db   3818 TATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATTGTGAGGGCAGCAAAGGAGGCC 3877

Qy    688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db   3878 AACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAA 3937

Qy    708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
Db   3938 GGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGA 3997

Qy    728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db   3998 CAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAG 4057
```

```
Qy     748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db    4058 GTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCAC 4117

Qy     768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db    4118 CGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGCAGAGTCAAG 4177

Qy     788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db    4178 GAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGT 4237

Qy     808 AlaGlnSer 810
Db    4238 GTCCAGGCT 4246
```

```
RESULT 13
US-09-306-417-1
; Sequence 1, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
;  APPLICANT: Heinrich-Pette-Institut
;  TITLE OF INVENTION: Retroviral Gene Transfer Vectors
;  FILE REFERENCE: P50491
;  CURRENT APPLICATION NUMBER: US/09/306,417
;  CURRENT FILING DATE: 1999-05-06
;  EARLIER APPLICATION NUMBER: DE 198 22 115
;  EARLIER FILING DATE: 1998-05-08
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 8630
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: proviral
;   OTHER INFORMATION: plasmid DNA
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(160)
;   OTHER INFORMATION: plasmid backbone (pUC)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (161)..(677)
;   OTHER INFORMATION: 5'-LTR
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (532)..(1219)
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: (1220)..(5062)
;   OTHER INFORMATION: m4 mdr-1 cDNA
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5215)..(5774)
;   OTHER INFORMATION: 3'-LTR
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5775)..(8630)
;   OTHER INFORMATION: plasmid backbone (pUC)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(8630)
;   OTHER INFORMATION: retroviral expression vector SFbeta71m4
US-09-306-417-1

Alignment Scores:
Pred. No.:              1.52e-229     Length:        8630
Score:                  2436.00       Matches:       466
Percent Similarity:     76.79%        Conservative:  166
Best Local Similarity:  56.62%        Mismatches:    177
Query Match:            59.72%        Indels:        14
DB:                     10            Gaps:          4
```

```
US-09-873-409-2 (1-812) x US-09-306-417-1 (1-8630)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db    2576 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 2635

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db    2636 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 2695

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db    2696 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 2755

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db    2756 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 2815

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db    2816 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 2875

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db    2876 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 2935

Qy     122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db    2936 CTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGTTTGTCTACAGTT 2995

Qy     142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db    2996 CGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 3055

Qy     162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178
Db    3056 GATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATGCAGACAGCAGGA 3115

Qy     179 ---------AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
Db    3116 AATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATTGATGCCTTGGAA 3175

Qy     196 ArgLysThrAsnSerLeuProLeuHisSerVal-----------------LysSerIle 209
Db    3176 ATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGTC 3235

Qy     210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228
Db    3236 CGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTGGATGAAAGTATA 3295

Qy     229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
Db    3296 CCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCCTTATTTTGTT 3355

Qy     249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db    3356 GTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTTGCAATAATATTT 3415

Qy     269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
Db    3416 TCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAACGACAGAATAGT 3475

Qy     288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db    3476 AACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATTACATTTTTCCTT 3535

Qy     308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db    3536 CAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTT 3595

Qy     328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
```

```
Db   3596 TTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAAAACACCACTGGA 3655

Qy    348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db   3656 GCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGG 3715

Qy    368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db   3716 CTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATC 3775

Qy    388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db   3776 TATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGA 3835

Qy    408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db   3836 GTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGT 3895

Qy    428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db   3896 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAG 3955

Qy    448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db   3956 GAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 4015

Qy    468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db   4016 AGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCC 4075

Qy    488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db   4076 TATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAG 4135

Qy    508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db   4136 GATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGT 4195

Qy    528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db   4196 TCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATT 4255

Qy    548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db   4256 GAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAA 4315

Qy    568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db   4316 GGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTG 4375

Qy    588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db   4376 CTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGT 4435

Qy    608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db   4436 GGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGG 4495

Qy    628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
Db   4496 AAAGTGCTGCTTGATGGCAAAGAAATTAAGCGACTGAATGTTCAGTGGCTCCGAGCACAC 4555

Qy    648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
Db   4556 CTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCC 4615

Qy    668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
Db   4616 TATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCC 4675

Qy    688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
Db   4676 AACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAA 4735

Qy    708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
Db   4736 GGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGA 4795

Qy    728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
Db   4796 CAGCCTCATATTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAG 4855

Qy    748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
Db   4856 GTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCAC 4915

Qy    768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
Db   4916 CGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGCAGAGTCAAG 4975

Qy    788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
Db   4976 GAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGT 5035

Qy    808 AlaGlnSer 810
Db   5036 GTCCAGGCT 5044
```

RESULT 14
US-09-306-417-2
; Sequence 2, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 8630
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: proviral
;   OTHER INFORMATION: plasmid DNA
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(8630)
;   OTHER INFORMATION: retroviral expression vector SFbeta91mSA1
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(160)
;   OTHER INFORMATION: plasmid backbone (pUC)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (161)..(677)
;   OTHER INFORMATION: 5'-LTR
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (532)..(1219)
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: (1220)..(5062)
;   OTHER INFORMATION: mSA1 mdr1 cDNA
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5215)..(5774)
;   OTHER INFORMATION: 3'-LTR
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5775)..(8630)
;   OTHER INFORMATION: plasmid backbone (pUC)

```
US-09-306-417-2

Alignment Scores:
Pred. No.:              1.52e-229      Length:         8630
Score:                  2436.00        Matches:        466
Percent Similarity:     76.79%         Conservative:   166
Best Local Similarity:  56.62%         Mismatches:     177
Query Match:            59.72%         Indels:         14
DB:                     10             Gaps:           4

US-09-873-409-2 (1-812) x US-09-306-417-2 (1-8630)

Qy        2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
Db     2576 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 2635

Qy       22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
Db     2636 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 2695

Qy       42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
Db     2696 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 2755

Qy       62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
Db     2756 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 2815

Qy       82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
Db     2816 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 2875

Qy      102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
Db     2876 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 2935

Qy      122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
Db     2936 CTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGTTTGTCTACAGTT 2995

Qy      142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
Db     2996 CGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAGAAAGGAAATCAT 3055

Qy      162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178
Db     3056 GATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATGCAGACAGCAGGA 3115

Qy      179 ---------AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 195
Db     3116 AATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATTGATGCCTTGGAA 3175

Qy      196 ArgLysThrAsnSerLeuProLeuHisSerVal------------------LysSerIle 209
Db     3176 ATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGTC 3235

Qy      210 Lys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228
Db     3236 CGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTGGATGAAAGTATA 3295

Qy      229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
Db     3296 CCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGGCCTTATTTTGTT 3355

Qy      249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
Db     3356 GTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTTGCAATAATATTT 3415

Qy      269 AlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAla 287
Db     3416 TCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAACGACAGAATAGT 3475

Qy      288 GluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMet 307
Db     3476 AACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATTACATTTTTCCTT 3535

Qy      308 GlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAla 327
Db     3536 CAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTT 3595

Qy      328 PheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGly 347
Db     3596 TTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAAAACACCACTGGA 3655

Qy      348 GlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArg 367
Db     3656 GCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGG 3715

Qy      368 IleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIle 387
Db     3716 CTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATC 3775

Qy      388 TyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGly 407
Db     3776 TATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGA 3835

Qy      408 MetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHis 427
Db     3836 GTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGT 3895

Qy      428 AlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArg 447
Db     3896 GCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAG 3955

Qy      448 GluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSer 467
Db     3956 GAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTG 4015

Qy      468 LysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAla 487
Db     4016 AGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCC 4075

Qy      488 TyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGlu 507
Db     4076 TATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAG 4135

Qy      508 GlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeu 527
Db     4136 GATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGT 4195

Qy      528 ValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeu 547
Db     4196 TCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATT 4255

Qy      548 GluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGlu 567
Db     4256 GAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAA 4315

Qy      568 GlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIle 587
Db     4316 GGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTG 4375

Qy      588 LeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSer 607
Db     4376 CTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGT 4435

Qy      608 GlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGly 627
Db     4436 GGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGG 4495

Qy      628 GlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln 647
Db     4496 AAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACAC 4555

Qy      648 IleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAla 667
```

```
Db   4556 CTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCC 4615

Qy    668 TyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 687
          ||||||||||||||||||||||        :::||||||      ||||||      |||
Db   4616 TATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCC 4675

Qy    688 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 707
          ||||||||:::|||||||||  ||||||  ||||||:::|||:::||||||   |||
Db   4676 AACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAA 4735

Qy    708 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 727
          |||   |||||||||||||||||||||||||:::||||||||||||||||||||::::::
Db   4736 GGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGA 4795

Qy    728 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 747
          :::|||   ||||||||||||||||||||||||||||||||||||     :::||||||||
Db   4796 CAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAG 4855

Qy    748 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 767
          |||||||||   ||||||||||||||||:||   ||||||||||:::|||:::   |||
Db   4856 GTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCAC 4915

Qy    768 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 787
          |||||||||   |||||||||||||||||||||||||||       ||||||::::::|||
Db   4916 CGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGCAGAGTCAAG 4975

Qy    788 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 807
          |||   ||||||||||||:::||||||        :::   |||||||||   :::|||:::
Db   4976 GAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGT 5035

Qy    808 AlaGlnSer 810
             |||:::
Db   5036 GTCCAGGCT 5044
```

RESULT 15
US-09-917-800A-1560
; Sequence 1560, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1560
; LENGTH: 3912
; TYPE: DNA
; ORGANISM: Rattus norvegicus
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_012690
US-09-917-800A-1560

Alignment Scores:			
Pred. No.:	1.31e-229	Length:	3912
Score:	2431.50	Matches:	470
Percent Similarity:	76.03%	Conservative:	155
Best Local Similarity:	57.18%	Mismatches:	184
Query Match:	59.61%	Indels:	13
DB:	10	Gaps:	2

US-09-873-409-2 (1-812) x US-09-917-800A-1560 (1-3912)

```
Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
          :::|||   |||||||||   |||||||||   |||:::   |||||||||
Db   1386 ATCGATGGGCAGGACATCCGGAACTTTAACGTCAGGTGTCTAAGGGAATTCATCGGCGTG 1445

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
          ||||||||||||||||||||||||   ||||||||||:::   ||||||:::|||||||||
Db   1446 GTGAGTCAAGAGCCGGTACTGTTCTCTACCACGATTGCTGAAAATATCCGCTATGGCCGT 1505

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
             :::||||||   :::|||:::::::::|||   :::||||||||||||||||||||||
Db   1506 GGGAATGTAACAATGGATGAGATTAAAAAAGCTGTCAAAGAGGCTAATGCCTATGACTTC 1565

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
          ||||||:::   |||   ||||||:::|||||||||||||::::::||||||||||:::|||
Db   1566 ATCATGAAACTGCCACAGAAATTTGACACCCTGGTTGGTGACAGAGGGGCGCAGCTGAGC 1625

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1626 GGGGGACAGAAACAGAGGATCGCCATTGCTCGTGCCTTGGTCCGCAACCCCAAGATCCTC 1685

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
          :::|||||||||||||||||||||||||||:::||||||::::::   ||||||||||||
Db   1686 CTGCTGGACGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCTGAGGTGCAGGCCGCT 1745

Qy    122 LeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIle 141
          |||:::||||||   :::||||||||||||||||||:::||||||||||||||||||:::
Db   1746 CTGGATAAGGCCAGAGAAGGCCGGACCACCATCGTGATAGCTCACCGACTGTCAACTGTC 1805

Qy    142 ArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHis 161
          |||:::|||||||||:::|||         :::||||||::::::   |||:::|||:::|||
Db   1806 CGGAATGCAGATGTCATCGCTGGGTTTGAGGATGGCGTCATCGTGGAGCAAGGAAGCCAC 1865

Qy    162 AlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln--------- 178
          :::||||||:::   |||   |||:::|||:::   ||||||      |||
Db   1866 AGTGAGCTGATAAAGAAGGAAGGGATCTACTTCAGACTTGTTAACATGCAGACATCAGGA 1925

Qy    179 ---------------------AspIleLysLysAlaAspGluGlnMetGluSerMet 190
                               :::::::::   :::||||||:::
Db   1926 AGCCAGATCCTGTCAGAAGAATTTGAAGTTGAGCTAAGTGATGAAAAGGCTGCTGGAGGT 1985

Qy    191 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 210
                            |||                       :::   ||||||:::|||
Db   1986 GTGGCCCCAAATGGCTGGAAAGCACGCATATTTAGGAATTCTACGAAGAAAAGTCTGAAA 2045

Qy    211 SerAsp------PheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeu 228
          |||         ::::::   :::         |||         :::   ::::::
Db   2046 AGTTCACGGGCGCATCAAAATAGGCTGGATGTGGAAACCAATGAACTTGATGCAAACGTG 2105

Qy    229 ProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValVal 248
          |||   ||||||   |||||:::|||:::||||||||||   ||||||||:::   |||
Db   2106 CCACCAGTGTCTTTTCTGAAGGTCTTAAGACTGAATAAAACAGAGTGGCCCTACTTTGTG 2165

Qy    249 LeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePhe 268
          :::|||||||||   ::::::   ||||||   :::   |||   ||||||||||||
Db   2166 GTGGGGACACTCTGTGCCATTGCCAACGGGGCCCTCCAGCCGGCATTCTCCATCATCCTG 2225

Qy    269 AlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGlu 288
          :::::::::|||   :::||||||         |||   |||         :::
Db   2226 TCAGAGATGATAGCTATCTTTGGCCCTGGGGATGACACAGTAAAGCAACAGAAGTGTAAC 2285
```

```
Qy      289 IleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGln 308
            ::::::|||::::::|||:::   |||||||||      |||   ::::::|||:::|||
Db     2286 ATGTTCTCGCTGGTCTTCTTGGGCCTAGGAGTCCACTCCTTCTTTACTTTCTTCCTTCAG 2345

Qy      309 GlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPhe 328
            |||      :::|||:::||||||||||||||||||   |||||||||   :::||||||
Db     2346 GGTTTCACATTCGGGAAAGCTGGCGAGATCCTCACCACAAGGCTCCGGTCCATGGCCTTC 2405

Qy      329 LysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGly 348
            ||||||||||||   ||||||::::::|||||||||:::   :::||||||||||||
Db     2406 AAAGCAATGCTAAGACAGGACATGAGCTGGTTTGACGATCATAAAAACAGTACTGGTGCC 2465

Qy      349 LeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIle 368
            |||:::|||   ||||||   |||   ||||||:::|||||||||||||||:::|||:::
Db     2466 CTCTCTACAAGACTCGCCACAGACGCTGCGCAGGTCCAAGGAGCCACAGGAACCAGGTTG 2525

Qy      369 GlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyr 388
               ::::::   ||||||      |||:::|||      :::||||||||||||||||||
Db     2526 GCTTTAATTGCACAGAACACAGCCAACCTTGGAACGGGTATTATTATATCATTTATTTAC 2585

Qy      389 GlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMet 408
            ||||||::::::|||   |||:::||||||:::   |||   :::||||||   |||:::
Db     2586 GGTTGGCAACTGACACTTCTGCTCTTATCAGTTGTTCCATTCATTGCTGTAGCGGGAATT 2645

Qy      409 IleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAla 428
            :::|||         :::   |||   |||   :::||||||:::|||::::::   |||
Db     2646 GTTGAAATGAAAATGTTGGCTGGCAACGCCAAGAGAGATAAAAAGGAGATGGAAGCTGCT 2705

Qy      429 GlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGlu 448
            |||||||||||||||||||||:::|||||||||||||||:::||||||||||||:::|||
Db     2706 GGAAAGATTGCAACAGAGGCAATAGAAAATATTCGGACTGTTGTATCCTTGACCCAAGAG 2765

Qy      449 LysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLys 468
            :::   ||||||   ||||||   |||   |||         :::||||||:::   :::
Db     2766 AGAAAATTTGAGTCAATGTATGTTGAAAAATTACACGGACCTTACAGGAATTCAGTGCGG 2825

Qy      469 LysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyr 488
            ||||||   |||   |||      ::::::   |||   ||||||:::||||||:::|||
Db     2826 AAGGCTCACATCTACGGCATCACTTTTAGCATCTCACAAGCATTCATGTACTTTTCTTAT 2885

Qy      489 AlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGly 508
            |||      ||||||||||||:::|||||||||      |||   |||      :::
Db     2886 GCTGGCTGCTTTCGATTTGGTTCTTACCTCATTGTGAATGGACACATGCGCTTCAAGGAT 2945

Qy      509 MetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuVal 528
            :::   :::||||||:::||||||      ||||||:::|||:::|||
Db     2946 GTCATCCTGGTGTTCTCAGCAATCGTGCTTGGTGCAGTGGCTCTAGGACATGCCAGCTCA 3005

Qy      529 LeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGlu 548
               ||||||:::|||:::|||||||||      ||||||:::||||||:::|||   |||
Db     3006 TTTGCTCCAGACTATGCAAAAGCCAAGCTGTCTGCAGCATACTTATTCAGTCTGTTTGAA 3065

Qy      549 LysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGly 568
            ::::::|||   |||||||||   |||:::||||||      ||||||      ||||||
Db     3066 AGACAACCTCTGATTGACAGCTACAGCAGAGAAGGAATGTGGCCGGATAAGTTTGAAGGA 3125

Qy      569 AsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeu 588
            ::::::   |||   ||||||   |||   ||||||   |||   :::|||   :::|||
Db     3126 AGCGTGACATTCAATGAAGTTGTGTTCAATTATCCCACCCGGGCCAATGTGCCAGTGCTT 3185

Qy      589 ArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGly 608
            :::||||||||||||   :::::::::|||:::|||:::|||   |||||||||||||||
Db     3186 CAGGGGCTGAGCCTCGAGGTGAAGAAGGGGCAGACCCTGGCCCTGGTGGGCAGTAGTGGC 3245

Qy      609 CysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGln 628
            |||||||||||||||   ||||||||||||:::|||   |||||||||:::   |||
Db     3246 TGCGGGAAGAGCACCGTGGTCCAGCTGCTCGAGCGCTTCTACGACCCCATGGCCGGAACA 3305

Qy      629 ValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIle 648
            ||||||   ||||||   :::||||||:::|||||||||||||||||||||:::|||:::
Db     3306 GTGCTCCTCGATGGTCAGGAAGCAAAGAAACTCAATGTCCAGTGGCTCCGAGCTCAACTT 3365

Qy      649 AlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyr 668
               ||||||   |||||||||:::||||||:::||||||||||||:::||||||||||||
Db     3366 GGCATTGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATCGCCAAGAACATCGCCTAC 3425

Qy      669 GlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsn 688
            |||||||||||||||||||||      |||||||||      ||||||      ||||||
Db     3426 GGAGACAACAGCCGTGTCGTGTCTCAGGATGAGATTGTGAGGGCGGCCAAGGAGGCCAAC 3485

Qy      689 IleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGly 708
            ||||||   |||||||||   ||||||:::||||||   |||:::||||||   ||||||
Db     3486 ATCCACCCCTTCATTGAGACACTGCCCCAAAAGTATGAAACAAGAGTAGGAGACAAGGGG 3545

Qy      709 AlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLys 728
               |||||||||||||||||||||||||||:::||||||||||||||||||:::::::::
Db     3546 ACACAGCTCTCTGGAGGCCAGAAACAGAGGATTGCTATCGCCCGAGCCCTCATCAGACAG 3605

Qy      729 ProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysVal 748
            |||::::::|||||||||||||||||||||||||||||||||   :::||||||||||||
Db     3606 CCTCGGGTCCTACTGCTGGATGAAGCCACGTCGGCTTTGGACACTGAGAGTGAAAAGGTC 3665

Qy      749 ValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArg 768
            ||||||   ||||||||||||||||||   ||||||||||||:::|||:::   ||||||
Db     3666 GTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACCTGCATTGTGATCGCGCACCGC 3725

Qy      769 LeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGlu 788
            ||||||   |||||||||||||||||||||||||||:::   |||||||||:::||||||
Db     3726 CTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATCGACAACGGCAAGGTCAAGGAG 3785

Qy      789 GlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 808
               ||||||||||||:::||||||      :::   |||||||||   :::||||||
Db     3786 CACGGCACCCACCAGCAGCTGCTGGCCCAGAAAGGCATCTATTTCTCCATGGTCAACATT 3845

Qy      809 GlnSer 810
            |||:::
Db     3846 CAAGCT 3851
```

Search completed: March 31, 2003, 16:00:18
Job time : 444.562 secs

```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 31, 2003, 13:52:37 ; Search time 239.871 Seconds
                                           (without alignments)
                                           2337.888 Million cell updates/sec

Title:          US-09-873-409-1
Perfect score:  3323
Sequence:       1 MLAEKGAHAELMAKRGLYYS...........QELLRNRDIYFKLVNAQSVQ 659

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters:      1148742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115424_19290/app_query.fasta_1.7544
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09873409_@CGN_1_1_960_@runat_27032003_115424_19290
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Published_Applications_NA:*
                1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
                2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
                3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
                4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
                5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
                6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
                7:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
                8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
                9:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
                10:  /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
                11:  /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
                12:  /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
                13:  /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
                14:  /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3323	100.0	2066	10	US-09-873-409-9	Sequence 9, Appli
2	3323	100.0	2856	10	US-09-873-409-10	Sequence 10, Appl
3	3323	100.0	3177	10	US-09-873-409-12	Sequence 12, Appl
4	3323	100.0	3621	10	US-09-873-409-14	Sequence 14, Appl
5	3323	100.0	3702	10	US-09-873-409-13	Sequence 13, Appl
6	1924.5	57.9	3924	10	US-09-880-107-2299	Sequence 2299, Ap
7	1882.5	56.7	3912	10	US-09-917-800A-1560	Sequence 1560, Ap
8	1882	56.6	4369	10	US-09-769-097-1	Sequence 1, Appli
9	1882	56.6	4425	10	US-09-769-097-3	Sequence 3, Appli
10	1872	56.3	3860	10	US-09-866-866A-1	Sequence 1, Appli
11	1872	56.3	3860	10	US-09-866-866A-3	Sequence 3, Appli
12	1872	56.3	4643	9	US-10-072-621-2	Sequence 2, Appli
13	1872	56.3	8630	10	US-09-306-417-1	Sequence 1, Appli
14	1872	56.3	8630	10	US-09-306-417-2	Sequence 2, Appli
15	1871	56.3	4788	10	US-09-866-866A-7	Sequence 7, Appli
16	1870	56.3	4317	9	US-10-044-671-1	Sequence 1, Appli
17	1844.5	55.5	4189	10	US-09-866-866A-5	Sequence 5, Appli
18	1813	54.6	4254	10	US-09-917-800A-1424	Sequence 1424, Ap
19	1516.5	45.6	5010	10	US-09-917-800A-483	Sequence 483, App
20	1254	37.7	4002	9	US-09-758-828-1	Sequence 1, Appli
21	1254	37.7	4002	9	US-09-758-828-3	Sequence 3, Appli
22	1251.5	37.7	3861	9	US-09-938-842A-2237	Sequence 2237, Ap
23	1235.5	37.2	2681	10	US-09-749-340-8	Sequence 8, Appli
24	1232	37.1	4175	10	US-09-749-340-3	Sequence 3, Appli
25	1226	36.9	2698	10	US-09-749-340-5	Sequence 5, Appli
26	1211.5	36.5	3512	10	US-09-749-340-7	Sequence 7, Appli
27	1098.5	33.1	3861	9	US-09-938-842A-263	Sequence 263, App
28	1053.5	31.7	4653	9	US-10-101-388-2	Sequence 2, Appli
29	1053.5	31.7	8036	9	US-10-101-388-1	Sequence 1, Appli
30	840	25.3	3792	9	US-09-882-694-10	Sequence 10, Appl
31	836.5	25.2	1810	10	US-09-749-340-4	Sequence 4, Appli
32	828	24.9	1940	10	US-09-873-409-16	Sequence 16, Appl
33	812.5	24.5	2021	10	US-09-873-409-15	Sequence 15, Appl
34	783	23.6	3999	9	US-09-882-694-9	Sequence 9, Appli
35	758.5	22.8	2298	9	US-10-156-239-18	Sequence 18, Appl
36	758.5	22.8	2298	10	US-09-795-693-18	Sequence 18, Appl
37	758.5	22.8	3408	9	US-10-156-239-16	Sequence 16, Appl
38	758.5	22.8	3408	10	US-09-795-693-16	Sequence 16, Appl
39	758.5	22.8	3512	9	US-10-072-621-1	Sequence 1, Appli
40	710.5	21.4	1749	10	US-09-815-242-6018	Sequence 6018, Ap
41	708.5	21.3	1749	10	US-09-815-242-9908	Sequence 9908, Ap
42	693.5	20.9	1812	10	US-09-815-242-4147	Sequence 4147, Ap
43	693	20.9	1761	9	US-10-260-877-25	Sequence 25, Appl
44	693	20.9	1764	10	US-09-815-242-6863	Sequence 6863, Ap
45	649	19.5	2247	9	US-09-870-759-40	Sequence 40, Appl

ALIGNMENTS

```
RESULT 1
US-09-873-409-9
; Sequence 9, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
;   LENGTH: 2066
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-409-9

Alignment Scores:
Pred. No.:              0                 Length:         2066
Score:                  3323.00           Matches:        659
Percent Similarity:     100.00%           Conservative:   0
Best Local Similarity:  100.00%           Mismatches:     0
Query Match:            100.00%           Indels:         0
DB:                     10                Gaps:           0
```

```
US-09-873-409-1 (1-659) x US-09-873-409-9 (1-2066)

Qy       1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      87 ATGCTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCA 146

Qy      21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     147 CTTGTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCT 206

Qy      41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     207 ACTGAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTC 266

Qy      61 IleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     267 ATTGACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTA 326

Qy      81 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     327 TTAAAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCT 386

Qy     101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     387 TCTGTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACC 446

Qy     121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     447 ATGTTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATA 506

Qy     141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     507 TTCGTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGC 566

Qy     161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     567 AGAGCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATAT 626

Qy     181 GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     627 CAGGATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTA 686

Qy     201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     687 GCCATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAA 746

Qy     221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     747 AATGCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACA 806

Qy     241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     807 TTCCTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCA 866

Qy     261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     867 ATGACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACT 926

Qy     281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     927 GAAGCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAA 986

Qy     301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     987 ATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATT 1046

Qy     321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1047 GGAAGCTGTTATGCATTCAGCCATGCCTTTATATATTTTGCCTATGCAGCAGGGTTTCGA 1106

Qy     341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1107 TTTGGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTT 1166

Qy     361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1167 ACTGCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATAT 1226

Qy     381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1227 TCCAAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATA 1286

Qy     401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1287 GACAGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGA 1346

Qy     421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1347 GAAGTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTC 1406

Qy     441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1407 AGTATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACT 1466

Qy     461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1467 TCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGT 1526

Qy     481 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1527 GTGGATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAA 1586

Qy     501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1587 GAGCCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGT 1646

Qy     521 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1647 GTGGTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATT 1706

Qy     541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1707 GAAGGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGC 1766

Qy     561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1767 GGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTG 1826

Qy     581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1827 TTGGATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTT 1886

Qy     601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1887 GATAAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAG 1946

Qy     621 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1947 AACGCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAA 2006

Qy     641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2007 GAGCTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 2063

RESULT 2
US-09-873-409-10
; Sequence 10, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
```

```
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
;   LENGTH: 2856
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-409-10

Alignment Scores:
Pred. No.:               0             Length:        2856
Score:                   3323.00       Matches:       659
Percent Similarity:      100.00%       Conservative:  0
Best Local Similarity:   100.00%       Mismatches:    0
Query Match:             100.00%       Indels:        0
DB:                      10            Gaps:          0

US-09-873-409-1 (1-659) x US-09-873-409-10 (1-2856)

Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    877 ATGCTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCA 936

Qy     21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    937 CTTGTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCT 996

Qy     41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    997 ACTGAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTC 1056

Qy     61 IleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1057 ATTGACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTA 1116

Qy     81 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1117 TTAAAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCT 1176

Qy    101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1177 TCTGTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACC 1236

Qy    121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1237 ATGTTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATA 1296

Qy    141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1297 TTCGTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGC 1356

Qy    161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1357 AGAGCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATAT 1416

Qy    181 GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1417 CAGGATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTA 1476

Qy    201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1477 GCCATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAA 1536

Qy    221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1537 AATGCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACA 1596

Qy    241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1597 TTCCTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCA 1656

Qy    261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1657 ATGACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACT 1716

Qy    281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1717 GAAGCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAA 1776

Qy    301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1777 ATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATT 1836

Qy    321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1837 GGAAGCTGTTATGCATTCAGCCATGCCTTTATATATTTGCCTATGCAGCAGGGTTTCGA 1896

Qy    341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1897 TTTGGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTT 1956

Qy    361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1957 ACTGCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATAT 2016

Qy    381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2017 TCCAAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATA 2076

Qy    401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2077 GACAGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGA 2136

Qy    421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2137 GAAGTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTC 2196

Qy    441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2197 AGTATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGAAAAAGCACT 2256

Qy    461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2257 TCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGT 2316

Qy    481 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2317 GTGGATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAA 2376

Qy    501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2377 GAGCCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGT 2436

Qy    521 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2437 GTGGTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATT 2496

Qy    541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2497 GAAGGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGC 2556

Qy    561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2557 GGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTG 2616

Qy    581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2617 TTGGATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTT 2676

Qy    601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2677 GATAAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAG 2736

Qy    621 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2737 AACGCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAA 2796
```

```
Qy     641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2797 GAGCTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 2853

RESULT 3
US-09-873-409-12
; Sequence 12, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
;   LENGTH: 3177
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: No. US20020037522A1e
;   LOCATION: (198)..(198)
;   OTHER INFORMATION: n at position 198 represents any nucleotide (A, T, C or G)
US-09-873-409-12

Alignment Scores:
Pred. No.:              0              Length:        3177
Score:                  3323.00        Matches:       659
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     10             Gaps:          0

US-09-873-409-1 (1-659) x US-09-873-409-12 (1-3177)

Qy       1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1198 ATGCTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCA 1257

Qy      21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1258 CTTGTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCT 1317

Qy      41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1318 ACTGAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTC 1377

Qy      61 IleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1378 ATTGACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTA 1437

Qy      81 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1438 TTAAAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCT 1497

Qy     101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1498 TCTGTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACC 1557

Qy     121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1558 ATGTTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATA 1617

Qy     141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1618 TTCGTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGC 1677

Qy     161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1678 AGAGCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATAT 1737

Qy     181 GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1738 CAGGATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTA 1797

Qy     201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1798 GCCATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAA 1857

Qy     221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1858 AATGCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACA 1917

Qy     241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1918 TTCCTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCA 1977

Qy     261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1978 ATGACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACT 2037

Qy     281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2038 GAAGCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAA 2097

Qy     301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2098 ATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATT 2157

Qy     321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2158 GGAAGCTGTTATGCATTCAGCCATGCCTTTATATATTTTGCCTATGCAGCAGGGTTTCGA 2217

Qy     341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2218 TTTGGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTT 2277

Qy     361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2278 ACTGCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATAT 2337

Qy     381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2338 TCCAAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATA 2397

Qy     401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2398 GACAGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGA 2457

Qy     421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2458 GAAGTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTC 2517

Qy     441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2518 AGTATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGAAAAAGCACT 2577

Qy     461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2578 TCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGT 2637

Qy     481 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2638 GTGGATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAA 2697

Qy     501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2698 GAGCCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGT 2757

Qy     521 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2758 GTGGTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATT 2817
```

```
Qy     541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2818 GAAGGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGC 2877

Qy     561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2878 GGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTG 2937

Qy     581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2938 TTGGATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTT 2997

Qy     601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2998 GATAAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAG 3057

Qy     621 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3058 AACGCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAA 3117

Qy     641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3118 GAGCTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3174

RESULT 4
US-09-873-409-14
; Sequence 14, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
;   LENGTH: 3621
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-409-14

Alignment Scores:
Pred. No.:              0              Length:        3621
Score:                  3323.00        Matches:       659
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     10             Gaps:          0

US-09-873-409-1 (1-659) x US-09-873-409-14 (1-3621)

Qy       1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1642 ATGCTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCA 1701

Qy      21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1702 CTTGTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCT 1761

Qy      41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1762 ACTGAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTC 1821

Qy      61 IleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1822 ATTGACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTA 1881

Qy      81 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1882 TTAAAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCT 1941

Qy     101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1942 TCTGTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACC 2001

Qy     121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2002 ATGTTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATA 2061

Qy     141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2062 TTCGTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGC 2121

Qy     161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2122 AGAGCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATAT 2181

Qy     181 GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2182 CAGGATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTA 2241

Qy     201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2242 GCCATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAA 2301

Qy     221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2302 AATGCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACA 2361

Qy     241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2362 TTCCTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCA 2421

Qy     261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2422 ATGACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACT 2481

Qy     281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2482 GAAGCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAA 2541

Qy     301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2542 ATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATT 2601

Qy     321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2602 GGAAGCTGTTATGCATTCAGCCATGCCTTTATATATTTTGCCTATGCAGCAGGGTTTCGA 2661

Qy     341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2662 TTTGGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTT 2721

Qy     361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2722 ACTGCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATAT 2781

Qy     381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2782 TCCAAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATA 2841

Qy     401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2842 GACAGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGA 2901

Qy     421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2902 GAAGTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTC 2961

Qy     441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2962 AGTATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACT 3021
```

```
Qy     461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3022 TCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGT 3081

Qy     481 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3082 GTGGATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAA 3141

Qy     501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3142 GAGCCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGT 3201

Qy     521 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3202 GTGGTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATT 3261

Qy     541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3262 GAAGGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGC 3321

Qy     561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3322 GGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTG 3381

Qy     581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3382 TTGGATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTT 3441

Qy     601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3442 GATAAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAG 3501

Qy     621 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3502 AACGCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAA 3561

Qy     641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3562 GAGCTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3618
```

```
RESULT 5
US-09-873-409-13
; Sequence 13, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
;   LENGTH: 3702
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: No. US20020037522A1e
;   LOCATION: (723)..(723)
;   OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-13

Alignment Scores:
Pred. No.:               0                Length:          3702
Score:                   3323.00          Matches:         659
Percent Similarity:      100.00%          Conservative:    0
Best Local Similarity:   100.00%          Mismatches:      0
Query Match:             100.00%          Indels:          0
DB:                      10               Gaps:            0

US-09-873-409-1 (1-659) x US-09-873-409-13 (1-3702)
```

```
Qy       1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1723 ATGCTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCA 1782

Qy      21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1783 CTTGTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCT 1842

Qy      41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1843 ACTGAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTC 1902

Qy      61 IleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1903 ATTGACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTA 1962

Qy      81 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1963 TTAAAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCT 2022

Qy     101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2023 TCTGTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACC 2082

Qy     121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2083 ATGTTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATA 2142

Qy     141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2143 TTCGTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGC 2202

Qy     161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2203 AGAGCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATAT 2262

Qy     181 GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2263 CAGGATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTA 2322

Qy     201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2323 GCCATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAA 2382

Qy     221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2383 AATGCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACA 2442

Qy     241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2443 TTCCTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCA 2502

Qy     261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2503 ATGACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACT 2562

Qy     281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2563 GAAGCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAA 2622

Qy     301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2623 ATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATT 2682

Qy     321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2683 GGAAGCTGTTATGCATTCAGCCATGCCTTTATATATTTGCCTATGCAGCAGGGTTTCGA 2742

Qy     341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2743 TTTGGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTT 2802
```

```
Qy     361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2803 ACTGCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATAT 2862

Qy     381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2863 TCCAAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATA 2922

Qy     401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2923 GACAGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGA 2982

Qy     421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2983 GAAGTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTC 3042

Qy     441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3043 AGTATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACT 3102

Qy     461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3103 TCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGT 3162

Qy     481 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3163 GTGGATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAA 3222

Qy     501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3223 GAGCCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGT 3282

Qy     521 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3283 GTGGTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATT 3342

Qy     541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3343 GAAGGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGC 3402

Qy     561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3403 GGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTG 3462

Qy     581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3463 TTGGATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTT 3522

Qy     601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3523 GATAAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAG 3582

Qy     621 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3583 AACGCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAA 3642

Qy     641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3643 GAGCTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3699
```

RESULT 6
US-09-880-107-2299
; Sequence 2299, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2299
; LENGTH: 3924
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 M23234
US-09-880-107-2299

```
Alignment Scores:
Pred. No.:              5.2e-195      Length:         3924
Score:                  1924.50       Matches:        372
Percent Similarity:     74.10%        Conservative:   123
Best Local Similarity:  55.69%        Mismatches:     162
Query Match:            57.91%        Indels:         11
DB:                     10            Gaps:           3

US-09-873-409-1 (1-659) x US-09-880-107-2299 (1-3924)

Qy       1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
             ::::::    |||:::|||:::|||:::|||||||||   |||   .|||:::|||:::
Db    1851 GTAATTGTGGAGCAAGGAAGCCACAGCGAACTGATGAAGAAGGAAGGGGTGTACTTCAAA 1910

Qy      21 LeuVal---------------------------MetSerGlnAspIleLysLysAlaAsp 31
             ||||||                             |||::::::   :::       |||
Db    1911 CTTGTCAACATGCAGACATCAGGAAGCCAGATCCAGTCAGAAGAATTTGAACTAAATGAT 1970

Qy      32 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 51
             |||:::          :::                  |||:::
Db    1971 GAAAAGGCTGCCACTAGAATGGCCCCAAATGGCTGGAAATCTCGCCTATTTAGGCATTCT 2030

Qy      52 SerValLysSerIleLys---SerAspPheIleAspLysAla---GluGluSerThrGln 69
           :::    |||::::::|||   |||            |||:::   :::       |||
Db    2031 ACTCAGAAAAACCTTAAAAATTCACAAATGTGTCAGAAGAGCCTTGATGTGGAAACCGAT 2090

Qy      70 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 89
                 |||    ::::::|||   ||||||    ||||||:::||||||||||||||
Db    2091 GGACTTGAAGCAAATGTGCCACCAGTGTCCTTTCTGAAGGTCCTGAAACTGAATAAAACA 2150

Qy      90 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 109
           |||||||||:::   |||:::|||||||||:::   ::::::   ||||||   :::   |||
Db    2151 GAATGGCCCTACTTTGTCGTGGGAACAGTATGTGCCATTGCCAATGGGGGGCTTCAGCCG 2210

Qy     110 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 129
              ||||||:::||||||::::::||||||   :::||||||      |||
Db    2211 GCATTTTCAGTCATATTCTCAGAGATCATAGCGATTTTTGGACCAGGCGATGATGCAGTG 2270

Qy     130 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 149
              :::               |||:::|||:::||||||:::    ||||||:::|||   |||
Db    2271 AAGCAGCAGAAGTGCAACATATTCTCTTTGATTTTCTTATTTCTGGGAATTATTTCTTTT 2330

Qy     150 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 169
              ::::::|||:::||||||      :::|||:::||||||||||||||||||   |||
Db    2331 TTTACTTTCTTCCTTCAGGGTTTCACGTTTGGGAAAGCTGGCGAGATCCTCACCAGAAGA 2390

Qy     170 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLys 189
           ||||||    :::||||||||||||||||||   ||||||::::::|||||||||:::
Db    2391 CTGCGGTCAATGGCTTTTAAAGCAATGCTAAGACAGGACATGAGCTGGTTTGATGACCAT 2450

Qy     190 GluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 209
           :::||||||||||||   |||:::|||   ||||||   |||   ||||||:::||||||
Db    2451 AAAAACAGTACTGGTGCACTTTCTACAAGACTTGCCACAGATGCTGCCCAAGTCCAAGGA 2510

Qy     210 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 229
           |||||||||:::|||:::      ::::::      ||||||      |||:::|||
Db    2511 GCCACAGGAACCAGGTTGGCTTTAATTGCACAGAATATAGCTAACCTTGGAACTGGTATT 2570

Qy     230 IleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVal 249
```

```
                |||||||||||||||||||||:::::|||   |||:::|||::::::   |||:::
Db     2571 ATCATATCATTTATCTACGGTTGGCAGTTAACCCTATTGCTATTAGCAGTTGTTCCAATT 2630

Qy      250 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 269
            :::||||||:::|||::::::|||        :::   |||    |||    :::||||||
Db     2631 ATTGCTGTGTCAGGAATTGTTGAAATGAAATTGTTGGCTGGAAATGCCAAAAGAGATAAA 2690

Qy      270 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIle 289
            :::||||||:::   |||||||||||||||||||||||||:::||||||||||||||||:::
Db     2691 AAAGAACTGGAAGCTGCTGGAAAGATTGCAACAGAGGCAATAGAAAATATTAGGACAGTT 2750

Qy      290 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGln 309
            ||||||||||||:::|||:::   ||||||   ||||||   |||   |||
Db     2751 GTGTCTTTGACCCAGGAAAGAAAATTTGAATCAATGTATGTTGAAAAATTGTATGGACCT 2810

Qy      310 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla 329
            :::||||||:::   :::||||||   |||   |||      ::::::   |||   |||
Db     2811 TACAGGAATTCTGTGCAGAAGGCACACATCTATGGAATTACTTTAGTATCTCACAAGCA 2870

Qy      330 PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGly 349
            |||:::||||||:::||||||      ||||||||||||||||||||||||      |||
Db     2871 TTTATGTATTTTTCCTATGCCGGTTGTTTTCGATTTGGTGCATATCTCATTGTGAATGGA 2930

Qy      350 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAla 369
               |||      :::   :::||||||:::||||||      :::||||||:::|||
Db     2931 CATATGCGCTTCAGAGATGTTATTCTGGTGTTTTCTGCAATTGTATTTGGTGCAGTGGCT 2990

Qy      370 IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHis 389
            :::|||               ||||||:::|||:::||||||||      |||||||||
Db     2991 CTAGGACATGCCAGTTCATTTGCTCCAGACTATGCTAAAGCTAAGCTGTCTGCAGCCCAC 3050

Qy      390 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLys 409
            ||||||   |||   |||::::::|||   ||||||||   |||:::||||||   |||
Db     3051 TTATTCATGCTGTTTGAAAGACAACCTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAG 3110

Qy      410 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 429
            ||||||      |||||||||:::   |||   ||||||   |||   ||||||   |||
Db     3111 CCTGATAAATTTGAAGGAAATATAACATTTAATGAAGTCGTGTTCAACTATCCCACCCGA 3170

Qy      430 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 449
               :::|||   :::|||||||||||||      :::::::::|||:::|||:::|||
Db     3171 GCAAACGTGCCAGTGCTTCAGGGGCTGAGCCTGGAGGTGAAGAAAGGCCAGACACTAGCC 3230

Qy      450 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr 469
               ||||||||||||||||||||||||||      ||||||||||||:::|||   |||
Db     3231 CTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACGGTGGTCCAGCTCCTGGAGCGGTTCTAC 3290

Qy      470 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 489
            ||||||:::   |||   ||||||   ||||||   :::||||||:::||||||||||||
Db     3291 GACCCCTTGGCGGGGACAGTGCTTCTCGATGGTCAAGAAGCAAAGAAACTCAATGTCCAG 3350

Qy      490 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 509
            ||||||||:::|||:::   ||||||   ||||||||||:::||||||:::||||||||
Db     3351 TGGCTCAGAGCTCAACTCGGAATCGTGTCTCAGGAGCCTATCCTATTTGACTGCAGCATT 3410

Qy      510 AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu 529
            ||||||||||||||||||||||||||||||||||||||         |||||||||
Db     3411 GCCGAGAATATTGCCTATGGAGACAACAGCCGGGTTGTATCACAGGATGAAATTGTGAGT 3470

Qy      530 AlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThr 549
            ||||||   |||||||||||||||   |||||||||   ||||||   ||||||   |||
Db     3471 GCAGCCAAAGCTGCCAACATACATCCTTTCATCGAGACGTTACCCCACAAATATGAAACA 3530

Qy      550 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAla 569
            :::||||||   ||||||   |||||||||||||||||||||||||||:::|||||||||
Db     3531 AGAGTGGGAGATAAGGGGACTCAGCTCTCAGGAGGTCAAAAACAGAGGATTGCTATTGCC 3590

Qy      570 ArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsp 589
            ||||||||:::::::::|||:::|||||||||||||||||||||||||||||||||||||
Db     3591 CGAGCCCTCATCAGACAACCTCAAATCCTCCTGTTGGATGAAGCTACATCAGCTCTGGAT 3650

Qy      590 AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys 609
               :::||||||||||||||||||   ||||||||||||||||||   ||||||||||||
```

```
Db     3651 ACTGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGC 3710

Qy      610 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 629
            :::|||:::   ||||||||||||   ||||||||||||||||||||||||||||
Db     3711 ATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAG 3770

Qy      630 AsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyr 649
            ||||||::::::||||||   |||||||||||:::||||||      :::   ||||||
Db     3771 AATGGGAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTAT 3830

Qy      650 PheLysLeuValAsnAlaGlnSer 657
            |||   :::|||:::   |||:::
Db     3831 TTTCAATGGTCAGTGTCCAGGCT 3854
```

RESULT 7
US-09-917-800A-1560
; Sequence 1560, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1560
;   LENGTH: 3912
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020119462A1 NM_012690
US-09-917-800A-1560

Alignment Scores:

Pred. No.:	1.56e-190	Length:	3912
Score:	1882.50	Matches:	363
Percent Similarity:	73.43%	Conservative:	129
Best Local Similarity:	54.18%	Mismatches:	165
Query Match:	56.65%	Indels:	13
DB:	10	Gaps:	2

US-09-873-409-1 (1-659) x US-09-917-800A-1560 (1-3912)

```
Qy        1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
               ::::::   |||:::|||:::|||:::||||||:::   |||   |||:::|||:::
Db     1842 GTCATCGTGGAGCAAGGAAGCCACAGTGAGCTGATAAAGAAGGAAGGATCTACTTCAGA 1901

Qy       21 LeuValMetSerGln--------------------------------AspIleLysLys 29
            ||||||   |||                                   ::::::::
```

```
Db   1902 CTTGTTAACATGCAGACATCAGGAAGCCAGATCCTGTCAGAAGAATTTGAAGTTGAGCTA 1961

Qy     30 AlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPro 49
Db   1962 AGTGATGAAAAGGCTGCTGGAGGTGTGGCCCCAAATGGCTGGAAAGCACGCATATTTAGG 2021

Qy     50 LeuHisSerValLysSerIleLysSerAsp------PheIleAspLysAlaGluGluSer 67
Db   2022 AATTCTACGAAGAAAAGTCTGAAAAGTTCACGGGCGCATCAAAATAGGCTGGATGTGGAA 2081

Qy     68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db   2082 ACCAATGAACTTGATGCAAACGTGCCACCAGTGTCTTTTCTGAAGGTCTTAAGACTGAAT 2141

Qy     88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db   2142 AAAACAGAGTGGCCCTACTTTGTGGTGGGGACACTCTGTGCCATTGCCAACGGGGCCCTC 2201

Qy    108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 127
Db   2202 CAGCCGGCATTCTCCATCATCCTGTCAGAGATGATAGCTATCTTTGGCCCTGGGGATGAC 2261

Qy    128 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 147
Db   2262 ACAGTAAAGCAACAGAAGTGTAACATGTTCTCGCTGGTCTTCTTGGGCCTAGGAGTCCAC 2321

Qy    148 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 167
Db   2322 TCCTTCTTTACTTTCTTCCTTCAGGGTTTCACATTCGGGAAAGCTGGCGAGATCCTCACC 2381

Qy    168 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 187
Db   2382 ACAAGGCTCCGGTCCATGGCCTTCAAAGCAATGCTAAGACAGGACATGAGCTGGTTTGAC 2441

Qy    188 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 207
Db   2442 GATCATAAAAACAGTACTGGTGCCCTCTCTACAAGACTCGCCACAGACGCTGCGCAGGTC 2501

Qy    208 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 227
Db   2502 CAAGGAGCCACAGGAACCAGGTTGGCTTTAATTGCACAGAACACAGCCAACCTTGGAACG 2561

Qy    228 SerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAla 247
Db   2562 GGTATTATTATATCATTTATTTACGGTTGGCAACTGACACTTCTGCTCTTATCAGTTGTT 2621

Qy    248 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 267
Db   2622 CCATTCATTGCTGTAGCGGGAATTGTTGAAATGAAAATGTTGGCTGGCAACGCCAAGAGA 2681

Qy    268 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 287
Db   2682 GATAAAAAGGAGATGGAAGCTGCTGGAAAGATTGCAACAGAGGCAATAGAAAATATTCGG 2741

Qy    288 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 307
Db   2742 ACTGTTGTATCCTTGACCCAAGAGAGAAAATTTGAGTCAATGTATGTTGAAAAATTACAC 2801

Qy    308 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 327
Db   2802 GGACCTTACAGGAATTCAGTGCGGAAGGCTCACATCTACGGCATCACTTTTAGCATCTCA 2861

Qy    328 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 347
Db   2862 CAAGCATTCATGTACTTTTCTTATGCTGGCTGCTTTCGATTTGGTTCTTACCTCATTGTG 2921

Qy    348 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 367
Db   2922 AATGGACACATGCGCTTCAAGGATGTCATCCTGGTGTTCTCAGCAATCGTGCTTGGTGCA 2981

Qy    368 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 387
Db   2982 GTGGCTCTAGGACATGCCAGCTCATTTGCTCCAGACTATGCAAAAGCCAAGCTGTCTGCA 3041

Qy    388 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 407
Db   3042 GCATACTTATTCAGTCTGTTTGAAAGACAACCTCTGATTGACAGCTACAGCAGAGAAGGA 3101

Qy    408 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 427
Db   3102 ATGTGGCCGGATAAGTTTGAAGGAAGCGTGACATTCAATGAAGTTGTGTTCAATTATCCC 3161

Qy    428 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 447
Db   3162 ACCCGGGCCAATGTGCCAGTGCTTCAGGGGCTGAGCCTCGAGGTGAAGAAGGGGCAGACC 3221

Qy    448 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 467
Db   3222 CTGGCCCTGGTGGGCAGTAGTGGCTGCGGGAAGAGCACCGTGGTCCAGCTGCTCGAGCGC 3281

Qy    468 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 487
Db   3282 TTCTACGACCCCATGGCCGGAACAGTGCTCCTCGATGGTCAGGAAGCAAAGAAACTCAAT 3341

Qy    488 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 507
Db   3342 GTCCAGTGGCTCCGAGCTCAACTTGGCATTGTGTCCCAGGAGCCCATCCTGTTTGACTGC 3401

Qy    508 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 527
Db   3402 AGCATCGCCAAGAACATCGCCTACGGAGACAACAGCCGTGTCGTGTCTCAGGATGAGATT 3461

Qy    528 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 547
Db   3462 GTGAGGGCGGCCAAGGAGGCCAACATCCACCCCTTCATTGAGACACTGCCCCAAAAGTAT 3521

Qy    548 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 567
Db   3522 GAAACAAGAGTAGGAGACAAGGGGACACAGCTCTCTGGAGGCCAGAAACAGAGGATTGCT 3581

Qy    568 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 587
Db   3582 ATCGCCCGAGCCCTCATCAGACAGCCTCGGGTCCTACTGCTGGATGAAGCCACGTCGGCT 3641

Qy    588 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 607
Db   3642 TTGGACACTGAGAGTGAAAAGGTCGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGC 3701

Qy    608 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 627
Db   3702 ACCTGCATTGTGATCGCGCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTG 3761

Qy    628 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 647
Db   3762 ATCGACAACGGCAAGGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCCCAGAAAGGC 3821

Qy    648 IleTyrPheLysLeuValAsnAlaGlnSer 657
Db   3822 ATCTATTTCTCCATGGTCAACATTCAAGCT 3851
```

RESULT 8
US-09-769-097-1
; Sequence 1, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: ENCODING RAT MDR1A AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP-50009-C2
; CURRENT APPLICATION NUMBER: US/09/769,097
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809

```
;  PRIOR FILING DATE: 1998-12-09
;  PRIOR APPLICATION NUMBER: 09/156,800
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: US99/20770
;  PRIOR FILING DATE: 1999-09-10
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 4369
;   TYPE: DNA
;   ORGANISM: RATTUS RATTUS
US-09-769-097-1

Alignment Scores:
Pred. No.:              2.1e-190     Length:        4369
Score:                  1882.00      Matches:       362
Percent Similarity:     74.37%       Conservative:  137
Best Local Similarity:  53.95%       Mismatches:    158
Query Match:            56.64%       Indels:        14
DB:                     10           Gaps:          5

US-09-873-409-1 (1-659) x US-09-769-097-1 (1-4369)

Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db   2140 GTCATTGTGGAGCAAGGAAATCATGATGAGCTCATGAGAGAGAAAGGAATTTACTTCAAA 2199

Qy     21 LeuValMetSerGln------------AspIleLysLysAlaAspGluGlnMetGluSer 36
Db   2200 CTTGTCATGACTCAGACAGCAGGAAATGAAATTGAATTAGGAAATGAAGCTTGTGAATCT 2259

Qy     37 ------------------MetThrTyrSerThrGluArgLysThrAsnSerLeu------ 48
Db   2260 AAAGACGGAATTGATAATGTGGACATGTCTTCAAAAGATTCGGGATCCAGTCTAATAAGA 2319

Qy     49 ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr 68
Db   2320 AGAAGATCAACTCGCAAAAGCATCCGTGGGCCACATGATCAAGACGGGGAACTTAGCACC 2379

Qy     69 ---GlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db   2380 AAAGAGGCTCTGGATGACGACGTACCTCCAGCTTCCTTTTGGCGGATCCTGAAGTTGAAT 2439

Qy     88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db   2440 TCAACTGAATGGCCTTATTTGTGGTTGGTGTATTTTGTGCCATAATAAATGGAGGCTTG 2499

Qy    108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 127
Db   2500 CAACCAGCATTCTCCATAATATTTCAAAGGTTGTAGGGGTTTTACAAAAAATGACACC 2559

Qy    128 ThrThrLeu---LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db   2560 CCTGAAATCCAGCGGCAGAACAGCAACTTGTTTTCTTTATTGTTCCTGATCCTTGGGATC 2619

Qy    147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db   2620 ATCTCTTTCATTACGTTTTTCCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTC 2679

Qy    167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
Db   2680 ACCAAGCGACTCCGATACATGGTCTTCAAATCCATGCTGAGACAGGACATAAGCTGGTTT 2739

Qy    187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db   2740 GATGACCCTAAAAACACCACAGGAGCGCTGACCACCAGGCTTGCCAATGACGCTGCTCAA 2799

Qy    207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db   2800 GTGAAAGGGGCTACAGGGTCTAGGCTTGCTGTTATTACCCAGAACATAGCAAATCTTGGG 2859

Qy    227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
Db   2860 ACAGGCATCATCATATCCCTGATCTACGGCTGGCAATTGACACTTTTACTCCTAGCAATT 2919

Qy    247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db   2920 GTTCCCATCATTGCTATAGCAGGAGTGGTTGAAATGAAAATGTTGTCTGGACAAGCGCTG 2979

Qy    267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db   2980 AAAGATAAGAAGGAACTAGAAGGTTCTGGGAAGATCGCTACAGAAGCAATTGAAAACTTT 3039

Qy    287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db   3040 CGCACTGTCGTCTCTTTGACTCGGGAGCAGAAGTTTGAAACTATGTATGCCCAGAGCTTG 3099

Qy    307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db   3100 CAGATACCATACAGAAATGCTTTGAAGAAAGCGCACGTCTTTGGGATCACTTTCTCCTTC 3159

Qy    327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
Db   3160 ACCCAGGCCATGATGTATTTCTCCTATGCTGCTTGTTTCCGGTTTGATGCCTACTTGGTG 3219

Qy    347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db   3220 GCACGAGAACTCATGACATTTGAAAATGTTCTGTTAGTATTCTCAGCTATTGTCTTTGGT 3279

Qy    367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db   3280 GCCATGGCAGTGGGGCAGGTCAGTTCATTCGCTCCTGACTACGCGAAAGCCAAAGTCTCG 3339

Qy    387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Db   3340 GCATCCCACATCATCAGGATCATTGAGAAAATCCCTGAGATTGACAGCTACAGCACGGAG 3399

Qy    407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
Db   3400 GGCTTGAAGCCTAATATGTTGGAAGGAAATGTGAAATTTAATGGAGTCATGTTCAACTAT 3459

Qy    427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db   3460 CCCACCCGACCCAACATCCCAGTGCTTCAGGGGCTGAGCCTAGAGGTGAAGAAGGGCAG 3519

Qy    447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
Db   3520 ACGCTGGCCCTCGTGGGCAGCAGTGGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTTGAG 3579

Qy    467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db   3580 CGCTTCTATGACCCCATGGCCGGAACAGTGTTTCTAGATGGCAAAGAAATAAAGCAACTC 3639

Qy    487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db   3640 AATGTCCAGTGGCTCCGCGCCCACCTGGGCATTGTGTCCCAGGAGCCCATCCTGTTTGAC 3699

Qy    507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db   3700 TGCAGCATCGCCGAGAACATTGCCTACGGAGACAACAGCCGTGTCGTGTCTCATAAGGAG 3759

Qy    527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db   3760 ATCGTGAAGGCAGCCAAGGAGGCCAACATCCACCAGTTCATCGACTCACTGCCTGAGAAA 3819

Qy    547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Db   3820 TACAACACCAGAGTGGGAGACAAAGGGACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATC 3879

Qy    567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Db   3880 GCCATCGCGCGCGCCCTCGTCAGACAGCCTCACATCTTACTTCTGGATGAAGCGACATCA 3939

Qy    587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
Db   3940 GCTCTGGATACGGAGAGTGAAAAGGTCGTCCAGGAAGCGCTGGACAAAGCCAGGAAGGC 3999
```

```
Qy    607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db   4000 CGCACCTGCATTGTGATCGCGCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTG 4059

Qy    627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
Db   4060 GTGATTCAGAACGGCCAGGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCCCAGAAA 4119

Qy    647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db   4120 GGCATCTATTTCTCGATGGTCAGTGTGCAGGCT 4152

RESULT 9
US-09-769-097-3
; Sequence 3, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
;  APPLICANT: Kimberly Anne Brun
;  APPLICANT:  Richard James Chenery
;  APPLICANT:  Harma Ellens
;  APPLICANT:  John Anthony Feild
;  APPLICANT:  Lin Yue
;  TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
;  TITLE OF INVENTION:  ENCODING RAT MDR1A AND SCREENING METHODS THEREOF
;  FILE REFERENCE: GP-50009-C2
;  CURRENT APPLICATION NUMBER: US/09/769,097
;  CURRENT FILING DATE:  2001-01-25
;  PRIOR APPLICATION NUMBER: 09/208,809
;  PRIOR FILING DATE: 1998-12-09
;  PRIOR APPLICATION NUMBER: 09/156,800
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: US99/20770
;  PRIOR FILING DATE: 1999-09-10
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 4425
;   TYPE: DNA
;   ORGANISM: RATTUS RATTUS
US-09-769-097-3

Alignment Scores:
Pred. No.:              2.14e-190      Length:          4425
Score:                  1882.00        Matches:         362
Percent Similarity:     74.37%         Conservative:    137
Best Local Similarity:  53.95%         Mismatches:      158
Query Match:            56.64%         Indels:          14
DB:                     10             Gaps:            5

US-09-873-409-1 (1-659) x US-09-769-097-3 (1-4425)

Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db   2140 GTCATTGTGGAGCAAGGAAATCATGATGAGCTCATGAGAGAGAAAGGAATTTACTTCAAA 2199

Qy     21 LeuValMetSerGln------------AspIleLysLysAlaAspGluGlnMetGluSer 36
Db   2200 CTTGTCATGACTCAGACAGCAGGAAATGAAATTGAATTAGGAAATGAAGCTTGTGAATCT 2259

Qy     37 ------------------MetThrTyrSerThrGluArgLysThrAsnSerLeu------ 48
Db   2260 AAAGAYGGAATTGATAATGTGGACATGTCTTCAAAAGATTCRGGATCCAGTCTAATAAGA 2319

Qy     49 ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr 68
Db   2320 AGAAGATCAACTCGCAAAAGCATCCGTGGGCCACATGATCAAGACGGGGAACTTAGCACC 2379

Qy     69 ---GlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db   2380 AAAGAGGCTCTGGATGACGACGTACCTCCAGCTTCCTTTTGGCGGATCCTGAAGTTGAAT 2439

Qy     88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db   2440 TCAACTGAATGGCCTTATTTGTGGTTGGTGTATTTGTGCCATAATAAATGGAGGCTTG 2499

Qy    108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 127
Db   2500 CAACCAGCATTCTCCATAATATTTTCAAAGGTTGTAGGGGTTTTTACAAAAAATGACACC 2559

Qy    128 ThrThrLeu---LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db   2560 CCTGAAATCCAGCGGCAGAACAGCAACTTGTTTTCTTTATTGTTCCTGATCCTTGGGATC 2619

Qy    147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db   2620 ATCTCTTTCATTACGTTTTTCCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTC 2679

Qy    167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
Db   2680 ACCAAGCGACTCCGATACATGGTCTTCAAATCCATGCTGAGACAGGACATAAGCTGGTTT 2739

Qy    187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db   2740 GATGACCCTAAAAACACCACAGGAGCGCTGACCACCAGGCTTGCCAATGACGCTGCTCAA 2799

Qy    207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db   2800 GTGAAAGGGGCTACAGGGTCTAGGCTTGCTGTTATTACCCAGAACATAGCAAATCTTGGG 2859

Qy    227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
Db   2860 ACAGGCATCATCATATCCCTGATCTACGGCTGGCAATTGACACTTTTACTCCTAGCAATT 2919

Qy    247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db   2920 GTTCCCATCATTGCTATAGCAGGAGTGGTTGAAATGAAAATGTTGTCTGGACAAGCGCTG 2979

Qy    267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db   2980 AAAGATAAGAAGGAACTAGAAGGTTCTGGGAAGATCGCTACAGAAGCAATTGAAAACTTT 3039

Qy    287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db   3040 CGCACTGTCGTCTCTTTGACTCGGGAGCAGAAGTTTGAAACTATGTATGCCCAGAGCTTG 3099

Qy    307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db   3100 CAGATACCATACAGAAATGCTTTGAAGAAAGCGCACGTCTTTGGGATCACTTTCTCCTTC 3159

Qy    327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
Db   3160 ACCCAGGCCATGATGTATTTCTCCTATGCTGCTTGTTTCCGGTTTGATGCCTACTTGGTG 3219

Qy    347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db   3220 GCACGAGAACTCATGACATTTGAAAATGTTCTGTTAGTATTCTCAGCTATTGTCTTTGGT 3279

Qy    367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db   3280 GCCATGGCAGTGGGGCAGGTCAGTTCATTCGCTCCTGACTACGCGAAAGCCAAAGTCTCG 3339

Qy    387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Db   3340 GCATCCCACATCATCAGGATCATTGAGAAAATCCCTGAGATTGACAGCTACAGCACGGAG 3399

Qy    407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
Db   3400 GGCTTGAAGCCTAATATGTTGGAAGGAAATGTGAAATTTAATGGAGTCATGTTCAACTAT 3459

Qy    427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db   3460 CCCACCCGACCCAACATCCCAGTGCTTCAGGGGCTGAGCCTAGAGGTGAAGAAAGGGCAG 3519

Qy    447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
```

```
Db   3520 ACGCTGGCCCTCGTGGGCAGCAGTGGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTTGAG 3579

Qy    467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
          |||   ||||||||:::   |||   |||     |||||   :::   |||:::|||
Db   3580 CGCTTCTATGACCCCATGGCCGGAACAGTGTTTCTAGATGGCAAAGAAATAAAGCAACTC 3639

Qy    487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
          ||||||||||||||||||:::   :::   |||||   |||||||||:::||||||:::
Db   3640 AATGTCCAGTGGCTCCGCGCCCACCTGGGCATTGTGTCCCAGGAGCCCATCCTGTTTGAC 3699

Qy    507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
          ||||||||||||||||||||||||||||||||||||||||||||||||         |||
Db   3700 TGCAGCATCGCCGAGAACATTGCCTACGGAGACAACAGCCGTGTCGTGTCTCATAAGGAG 3759

Qy    527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
          |||   :::||||||      |||||||||||   ||||||:::   |||||||||||
Db   3760 ATCGTGAAGGCAGCCAAGGAGGCCAACATCCACCAGTTCATCGACTCACTGCCTGAGAAA 3819

Qy    547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
          |||||||||:::||||||   ||||||   ||||||||||||||||||||||||:::
Db   3820 TACAACACCAGAGTGGGAGACAAAGGGACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATC 3879

Qy    567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
          ||||||||||||||||||:::::::::|||   |||||||||||||||||||||||||||
Db   3880 GCCATCGCGCGCGCCCTCGTCAGACAGCCTCACATCTTACTTCTGGATGAAGCGACATCA 3939

Qy    587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
          ||||||||     :::||||||||||||   ||||||||||||||||||||     |||
Db   3940 GCTCTGGATACGGAGAGTGAAAAGGTCGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGC 3999

Qy    607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
          |||||||||:::|||:::   ||||||||||||   ||||||||||||||||||||||||
Db   4000 CGCACCTGCATTGTGATCGCGCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTG 4059

Qy    627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
          |||:::   ||||||::::::||||||   ||||||||||||:::||||||      :::
Db   4060 GTGATTCAGAACGGCCAGGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCCCAGAAA 4119

Qy    647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
             |||||||||   :::|||:::   |||:::
Db   4120 GGCATCTATTTCTCGATGGTCAGTGTGCAGGCT 4152
```

RESULT 10
US-09-866-866A-1
; Sequence 1, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3860
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-866A-1

Alignment Scores:

Pred. No.:	2.02e-189	Length:	3860
Score:	1872.00	Matches:	356
Percent Similarity:	74.22%	Conservative:	142
Best Local Similarity:	53.06%	Mismatches:	159
Query Match:	56.33%	Indels:	14
DB:	10	Gaps:	4

```
US-09-873-409-1 (1-659) x US-09-866-866A-1 (1-3860)

Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
          ::::::   ||||||||   |||   |||||||||   ::::::|||:::|||:::
Db   1813 GTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAA 1872

Qy     21 LeuValMetSerGln-----------------AspIleLysLysAlaAspGluGlnMet 34
          ||||||   |||                 :::::::::   |||   ::::::
Db   1873 CTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCC 1932

Qy     35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
          :::|||      :::   |||      :::|||                   :::
Db   1933 AAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGA 1992

Qy     54 --------------LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
                   :::|||::::::   |||           |||:::         :::
Db   1993 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACC 2052

Qy     68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
          ::::::   :::      |||:::|||   ||||||      :::|||:::|||||||||
Db   2053 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAAT 2112

Qy     88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
                |||||||||:::   |||:::|||      :::::::::||||||   :::
Db   2113 TTAACTGAATGGCCTTATTTGTTGTTGGTGTATTTGTGCCATTATAAATGGAGGCCTG 2172

Qy    108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
             |||   |||:::|||||||||:::|||||||||   :::|||         :::|||
Db   2173 CAACCAGCATTTGCAATAATATTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGAT 2232

Qy    127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
             |||   :::   ::::::   ::::::|||::::::|||:::   |||||||:::
Db   2233 CCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATT 2292

Qy    147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
          |||   |||:::::::::|||:::||||||      :::|||:::|||||||||||||||
Db   2293 ATTTCTTTTATTACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTC 2352

Qy    167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
          |||   |||||||||::::::   |||::::::||||||   ||||||::::::||||||
Db   2353 ACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 2412

Qy    187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
          |||:::   :::|||:::||||||   |||||||||   ||||||   |||   ||||||
Db   2413 GATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAA 2472

Qy    207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
          ::::::||||||   |||||||||:::   |||:::|||||||||      |||:::|||
Db   2473 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGG 2532

Qy    227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
             :::||||||||||||||||||||||||||||::::::|||   |||:::|||:::|||
Db   2533 ACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATT 2592

Qy    247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
             |||::::::|||:::   |||::::::|||         ::::::|||   |||
Db   2593 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 2652

Qy    267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
          |||||||||:::||||||:::   |||||||||||||||||||||||||:::||||||
Db   2653 AAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 2712

Qy    287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
          ||||||:::|||||||||||||:::|||:::   ||||||   ||||||   :::   |||
Db   2713 CGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTG 2772

Qy    307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
          |||         :::|||||||:::   :::||||||   |||   |||      ::::::|||
Db   2773 CAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTC 2832

Qy    327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
```

```
Db   2833 ACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTG 2892

Qy    347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db   2893 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGT 2952

Qy    367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db   2953 GCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCA 3012

Qy    387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Db   3013 GCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAA 3072

Qy    407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
Db   3073 GGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTAT 3132

Qy    427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db   3133 CCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAG 3192

Qy    447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
Db   3193 ACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAG 3252

Qy    467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db   3253 CGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTG 3312

Qy    487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db   3313 AATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGAC 3372

Qy    507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db   3373 TGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAG 3432

Qy    527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db   3433 ATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAA 3492

Qy    547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Db   3493 TATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATT 3552

Qy    567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Db   3553 GCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCA 3612

Qy    587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
Db   3613 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 3672

Qy    607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db   3673 CGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTG 3732

Qy    627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
Db   3733 GTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAA 3792

Qy    647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db   3793 GGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3825
```

RESULT 11
US-09-866-866A-3
; Sequence 3, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
;   LENGTH: 3860
;   TYPE: DNA
;   ORGANISM: homo sapiens
US-09-866-866A-3

```
Alignment Scores:
Pred. No.:              2.02e-189     Length:         3860
Score:                  1872.00       Matches:        356
Percent Similarity:     74.22%        Conservative:   142
Best Local Similarity:  53.06%        Mismatches:     159
Query Match:            56.33%        Indels:         14
DB:                     10            Gaps:           4

US-09-873-409-1 (1-659) x US-09-866-866A-3 (1-3860)

Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db   1813 GTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAA 1872

Qy     21 LeuValMetSerGln-----------------AspIleLysLysAlaAspGluGlnMet 34
Db   1873 CTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCC 1932

Qy     35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
Db   1933 AAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGA 1992

Qy     54 --------------LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
Db   1993 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACC 2052

Qy     68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db   2053 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAAT 2112

Qy     88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db   2113 TTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTG 2172

Qy    108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
Db   2173 CAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGAT 2232

Qy    127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db   2233 CCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATT 2292

Qy    147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db   2293 ATTTCTTTTATTACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTC 2352

Qy    167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
Db   2353 ACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 2412

Qy    187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db   2413 GATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAA 2472
```

```
Qy    207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db   2473 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGG 2532

Qy    227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
Db   2533 ACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATT 2592

Qy    247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db   2593 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 2652

Qy    267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db   2653 AAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 2712

Qy    287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db   2713 CGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTG 2772

Qy    307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db   2773 CAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTC 2832

Qy    327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
Db   2833 ACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTG 2892

Qy    347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db   2893 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGT 2952

Qy    367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db   2953 GCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCA 3012

Qy    387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Db   3013 GCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAA 3072

Qy    407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
Db   3073 GGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTAT 3132

Qy    427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db   3133 CCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAG 3192

Qy    447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
Db   3193 ACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAG 3252

Qy    467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db   3253 CGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTG 3312

Qy    487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db   3313 AATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGAC 3372

Qy    507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db   3373 TGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAG 3432

Qy    527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db   3433 ATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAA 3492

Qy    547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Db   3493 TATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATT 3552

Qy    567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Db   3553 GCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCA 3612

Qy    587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
Db   3613 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 3672

Qy    607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db   3673 CGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTG 3732

Qy    627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
Db   3733 GTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAA 3792

Qy    647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db   3793 GGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3825
```

RESULT 12
US-10-072-621-2
; Sequence 2, Application US/10072621
; Patent No. US20020169137A1
; GENERAL INFORMATION:
; APPLICANT: Reiner, Peter B.
; APPLICANT: Connop, Bruce P.
; APPLICANT: Pollard, Michelle
; TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION
; TITLE OF INVENTION: BY MODIFICATION OF ABC TRANSPORTER EXPRESSION OR ACTIVITY
; FILE REFERENCE: 100103.402
; CURRENT APPLICATION NUMBER: US/10/072,621
; CURRENT FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 4643
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-072-621-2

Alignment Scores:

Pred. No.:	2.68e-189	Length:	4643
Score:	1872.00	Matches:	356
Percent Similarity:	74.22%	Conservative:	142
Best Local Similarity:	53.06%	Mismatches:	159
Query Match:	56.33%	Indels:	14
DB:	9	Gaps:	4

US-09-873-409-1 (1-659) x US-10-072-621-2 (1-4643)

```
Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db   2234 GTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAA 2293

Qy     21 LeuValMetSerGln------------------AspIleLysLysAlaAspGluGlnMet 34
Db   2294 CTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCC 2353

Qy     35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
Db   2354 AAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGA 2413

Qy     54 ---------------LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
Db   2414 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACC 2473

Qy     68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db   2474 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAAT 2533

Qy     88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db   2534 TTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTG 2593
```

```
Qy      108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
             |||   |||:::||||||||:::||||||||    :::|||        :::|||
Db     2594 CAACCAGCATTTGCAATAATATTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGAT 2653

Qy      127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
                     |||   :::   ::::::   ::::::|||::::::|||:::   ||||||:::
Db     2654 CCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATT 2713

Qy      147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
            |||   |||:::::::::|||:::||||||         :::|||:::||||||||||||||
Db     2714 ATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTC 2773

Qy      167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
            |||   |||||||||::::::   |||::::::|||||||   ||||||::::::||||||
Db     2774 ACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 2833

Qy      187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
            |||:::   :::|||:::||||||   |||||||||   ||||||   |||   ||||||
Db     2834 GATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAA 2893

Qy      207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
            ::::::||||||   |||||||||:::   |||:::|||||||||      |||:::|||
Db     2894 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGG 2953

Qy      227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
               :::||||||||||||||||||||||||||||:::::::|||   |||:::|||:::|||
Db     2954 ACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATT 3013

Qy      247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
               |||::::::|||:::   |||::::::|||         ::::::|||   |||
Db     3014 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 3073

Qy      267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
            ||||||||:::   |||||||||||||||||||||||||||||||||:::||||||
Db     3074 AAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 3133

Qy      287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
            ||||||:::||||||||||||:::|||:::   ||||||   ||||||   :::   |||
Db     3134 CGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTG 3193

Qy      307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
            |||      :::||||||:::   :::||||||   |||   |||      ::::::|||
Db     3194 CAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTC 3253

Qy      327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
            :::   |||   :::||||||:::||||||      |||||||||||||||||||||:::
Db     3254 ACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTG 3313

Qy      347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
                     |||:::   |||   :::   :::||||||:::|||:::   :::|||
Db     3314 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGT 3373

Qy      367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
            ||||||||:::|||:::                  ||||||:::|||:::||||||||
Db     3374 GCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCA 3433

Qy      387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
            |||||||||:::         ::::::||||||   |||   |||||||||   |||   |||
Db     3434 GCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAA 3493

Qy      407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
            |||         |||:::|||   |||||||||:::   |||   ||||||   |||   |||
Db     3494 GGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTAT 3553

Qy      427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
            |||   |||||||||:::   :::|||:::|||||||||      :::::::::|||:::
Db     3554 CCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAG 3613

Qy      447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
            |||:::|||   |||||||||||||||||||||||||||      ||||||||||||:::
Db     3614 ACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAG 3673

Qy      467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
            |||   ||||||||||:::   |||:::||||||   ||||||   :::   |||   |||
Db     3674 CGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTG 3733

Qy      487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
            ||||||||||||||||||:::   :::      ||||||   |||||||||:::||||||:::
Db     3734 AATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGAC 3793

Qy      507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
            ||||||||||||||||||||||||||||||||||||||||||||||||         :::|||
Db     3794 TGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAG 3853

Qy      527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
            |||      ||||||      ||||||||||||:::|||||||||   ||||||   |||
Db     3854 ATTGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAA 3913

Qy      547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
            |||:::|||:::||||||   ||||||   |||||||||||||||||||||||||||:::
Db     3914 TATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATT 3973

Qy      567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
            |||||||||||||||||||:::::::::|||   ||||||||||||||||||||||||||||
Db     3974 GCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCA 4033

Qy      587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
            |||||||||   :::||||||||||||||||||   ||||||||||||||||||   |||
Db     4034 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 4093

Qy      607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
            |||||||||:::|||:::   ||||||||||||   ||||||||||||||||||||||||
Db     4094 CGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTG 4153

Qy      627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
            |||      ||||||::::::||||||   ||||||||||||:::||||||      :::
Db     4154 GTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAA 4213

Qy      647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
               |||||||||   :::|||:::   |||:::
Db     4214 GGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 4246
```

RESULT 13
US-09-306-417-1
; Sequence 1, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 8630
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: proviral
;   OTHER INFORMATION: plasmid DNA
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(160)
;   OTHER INFORMATION: plasmid backbone (pUC)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (161)..(677)
;   OTHER INFORMATION: 5'-LTR
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (532)..(1219)
;   FEATURE:

```
;    NAME/KEY: mat_peptide
;    LOCATION: (1220)..(5062)
;    OTHER INFORMATION: m4 mdr-1 cDNA
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (5215)..(5774)
;    OTHER INFORMATION: 3'-LTR
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (5775)..(8630)
;    OTHER INFORMATION: plasmid backbone (pUC)
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (1)..(8630)
;    OTHER INFORMATION: retroviral expression vector SFbeta71m4
US-09-306-417-1

Alignment Scores:
Pred. No.:               6.95e-189     Length:          8630
Score:                   1872.00       Matches:         356
Percent Similarity:      74.22%        Conservative:    142
Best Local Similarity:   53.06%        Mismatches:      159
Query Match:             56.33%        Indels:          14
DB:                      10            Gaps:            4

US-09-873-409-1 (1-659) x US-09-306-417-1 (1-8630)

Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db   3032 GTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAA 3091

Qy     21 LeuValMetSerGln------------------AspIleLysLysAlaAspGluGlnMet 34
Db   3092 CTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCC 3151

Qy     35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
Db   3152 AAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGA 3211

Qy     54 ---------------LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
Db   3212 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACC 3271

Qy     68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db   3272 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAAT 3331

Qy     88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db   3332 TTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTG 3391

Qy    108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
Db   3392 CAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGAT 3451

Qy    127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db   3452 CCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATT 3511

Qy    147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db   3512 ATTTCTTTTATTACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTC 3571

Qy    167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
Db   3572 ACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 3631

Qy    187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db   3632 GATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAA 3691

Qy    207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db   3692 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGG 3751

Qy    227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
Db   3752 ACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATT 3811

Qy    247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db   3812 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 3871

Qy    267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db   3872 AAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 3931

Qy    287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db   3932 CGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTG 3991

Qy    307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db   3992 CAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTC 4051

Qy    327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
Db   4052 ACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTG 4111

Qy    347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db   4112 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGT 4171

Qy    367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db   4172 GCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCA 4231

Qy    387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Db   4232 GCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAA 4291

Qy    407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
Db   4292 GGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTAT 4351

Qy    427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db   4352 CCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAG 4411

Qy    447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
Db   4412 ACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAG 4471

Qy    467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db   4472 CGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATTAAGCGACTG 4531

Qy    487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db   4532 AATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGAC 4591

Qy    507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db   4592 TGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAG 4651

Qy    527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db   4652 ATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAA 4711

Qy    547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Db   4712 TATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATT 4771

Qy    567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Db   4772 GCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCA 4831
```

```
Qy     587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
           ||||||||   :::||||||||||||||||||   ||||||||||||||||||   |||
Db    4832 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 4891

Qy     607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
           ||||||||:::|||:::   ||||||||||||   |||||||||||||||||||||||
Db    4892 CGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTG 4951

Qy     627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
           |||      ||||||::::::||||||   ||||||||||||:::||||||         :::
Db    4952 GTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAA 5011

Qy     647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
              |||||||||   :::|||:::   |||:::
Db    5012 GGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 5044
```

RESULT 14
US-09-306-417-2
; Sequence 2, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: proviral
; OTHER INFORMATION: plasmid DNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(8630)
; OTHER INFORMATION: retroviral expression vector SFbeta91mSA1
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(160)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (161)..(677)
; OTHER INFORMATION: 5'-LTR
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (532)..(1219)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (1220)..(5062)
; OTHER INFORMATION: mSA1 mdr1 cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5215)..(5774)
; OTHER INFORMATION: 3'-LTR
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5775)..(8630)
; OTHER INFORMATION: plasmid backbone (pUC)
US-09-306-417-2

Alignment Scores:

Pred. No.:	6.95e-189	Length:	8630
Score:	1872.00	Matches:	356
Percent Similarity:	74.22%	Conservative:	142
Best Local Similarity:	53.06%	Mismatches:	159
Query Match:	56.33%	Indels:	14

DB: 10 Gaps: 4

US-09-873-409-1 (1-659) x US-09-306-417-2 (1-8630)

```
Qy       1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
           ::::::   ||||||||   |||   |||||||||   ::::::|||:::|||:::
Db    3032 GTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAA 3091

Qy      21 LeuValMetSerGln------------------AspIleLysLysAlaAspGluGlnMet 34
           ||||||   |||                     :::::::::   |||   ::::::
Db    3092 CTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCC 3151

Qy      35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
           :::|||      :::   |||      :::|||                  :::
Db    3152 AAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGA 3211

Qy      54 --------------LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
                         :::|||::::::   |||         |||:::         :::
Db    3212 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACC 3271

Qy      68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
              ::::::   :::   |||:::|||   ||||||      :::|||:::|||||||||
Db    3272 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAAT 3331

Qy      88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
                 |||||||||:::   |||:::|||      :::::::::||||||   :::
Db    3332 TTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTGTGCCATTATAAATGGAGGCCTG 3391

Qy     108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
              |||   |||:::|||||||||:::|||||||||      :::|||         :::|||
Db    3392 CAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTACAAGAATTGATGAT 3451

Qy     127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
              |||   :::      ::::::|||::::::|||:::      ||||||:::
Db    3452 CCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATT 3511

Qy     147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
           |||   |||:::::::::|||:::|||      :::|||:::|||||||||||||||
Db    3512 ATTTCTTTTATTACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTC 3571

Qy     167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
           |||   |||||||||::::::   |||::::::||||||   ||||||::::::||||||
Db    3572 ACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 3631

Qy     187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
           |||:::   :::|||:::||||||   |||||||||   ||||||   |||   ||||||
Db    3632 GATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAA 3691

Qy     207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
           ::::::||||||   |||||||||:::   |||:::|||||||||      |||:::|||
Db    3692 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGG 3751

Qy     227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
              :::||||||||||||||||||||||||||||::::::|||   |||:::|||:::|||
Db    3752 ACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATT 3811

Qy     247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
              |||::::::|||:::   |||::::::|||         ::::::|||   |||
Db    3812 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 3871

Qy     267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
           ||||||:::||||||:::   ||||||||||||||||||||||||:::||||||
Db    3872 AAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 3931

Qy     287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
           ||||||:::||||||||||||:::|||:::   ||||||   ||||||   :::   |||
Db    3932 CGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTG 3991

Qy     307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
           |||      :::||||||:::   :::||||||   |||   |||      ::::::|||
Db    3992 CAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTC 4051

Qy     327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
```

```
              :::    |||    :::||||||:::||||||          ||||||||||||||||||||:::
Db     4052 ACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTG 4111

Qy      347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
                         |||:::   |||    :::    :::||||||:::|||:::    :::|||
Db     4112 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGT 4171

Qy      367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
            |||||||||:::|||:::                      ||||||:::|||:::|||||||||
Db     4172 GCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCA 4231

Qy      387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
            |||||||||:::          ::::::||||||   |||   |||||||||   |||   |||
Db     4232 GCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAA 4291

Qy      407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
            |||          |||:::|||   |||||||||:::   |||   ||||||   |||   |||
Db     4292 GGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTAT 4351

Qy      427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
            |||   ||||||||:::   :::|||:::|||||||||||||       :::::::::|||:::
Db     4352 CCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAG 4411

Qy      447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
            |||:::|||   ||||||||||||||||||||||||||||||||||   |||||||||:::
Db     4412 ACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAG 4471

Qy      467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
            |||   |||||||||:::   |||:::||||||   ||||||    :::   |||   |||
Db     4472 CGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTG 4531

Qy      487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
            ||||||||||||||||||:::    :::    ||||||   |||||||||:::||||||:::
Db     4532 AATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGAC 4591

Qy      507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
            |||||||||||||||||||||||||||||||||||||||||||||||||      :::|||
Db     4592 TGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAG 4651

Qy      527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
            |||         ||||||            ||||||||||:::||||||||||   ||||||   |||
Db     4652 ATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAA 4711

Qy      547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
            |||:::|||:::||||||   ||||||   ||||||||||||||||||||||||||:::
Db     4712 TATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATT 4771

Qy      567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
            ||||||||||||||||||:::::::::|||   ||||||||||||||||||||||||||||
Db     4772 GCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTGCTTTTGGATGAAGCCACGTCA 4831

Qy      587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
            ||||||||      :::||||||||||||||||||   ||||||||||||||||||   |||
Db     4832 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 4891

Qy      607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
            |||||||||:::|||:::      ||||||||||||   ||||||||||||||||||||||||
Db     4892 CGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTG 4951

Qy      627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
            |||      ||||||:::::::||||||   ||||||||||||:::||||||       :::
Db     4952 GTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAA 5011

Qy      647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
                ||||||||||    :::|||:::    |||:::
Db     5012 GGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 5044
```

RESULT 15
US-09-866-866A-7
; Sequence 7, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
;   LENGTH: 4788
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-866-866A-7

Alignment Scores:

Pred. No.:	3.6e-189	Length:	4788
Score:	1871.00	Matches:	352
Percent Similarity:	73.70%	Conservative:	144
Best Local Similarity:	52.30%	Mismatches:	159
Query Match:	56.30%	Indels:	18
DB:	10	Gaps:	4

US-09-873-409-1 (1-659) x US-09-866-866A-7 (1-4788)

```
Qy        1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
             ::::::   |||:::|||   |||   |||||||||    ::::::|||:::|||:::
Db     1801 GTCATTGTGGAGCAAGGAAATCATGATGAGCTCATGAGAGAAAAGGGCATTTACTTCAAA 1860

Qy       21 LeuValMetSerGln------------------------------AspIleLysLysAla 30
            |||||||||:::|||                               :::      |||:::
Db     1861 CTTGTCATGACACAGACAGCAGGAAATGAAATTGAATTAGGAAATGAAGCTTGTAAATCT 1920

Qy       31 AspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeu 50
               ::::::::::::::::::         |||::::::          :::||||||
Db     1921 AAGGATGAAATTGATAATTTAGACATGTCTTCAAAAGATTCAGGATCCAGTCTAATAAGA 1980

Qy       51 HisSerValLysSerIleLysSerAspPheIle---------------AspLysAlaGlu 65
                  :::|||    :::                              |||:::
Db     1981 AGA------AGATCAACTCGCAAAAGCATCTGTGGACCACATGACCAAGACAGGAAGCTT 2034

Qy       66 GluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLys 85
               :::   ::::::   :::         :::|||      |||      :::|||||||||
Db     2035 AGTACCAAAGAGGCCCTGGATGAAGATGTACCTCCAGCTTCCTTTTGGCGGATCCTGAAG 2094

Qy       86 LeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGly 105
            ||||||      |||||||||:::   |||:::|||         :::::::::||||||
Db     2095 TTGAATTCAACTGAATGGCCTTATTTGTGGTTGGTATATTCTGTGCCATAATAAATGGA 2154

Qy      106 ThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn--- 124
               :::   |||   ||||||:::||||||:::|||::::::      :::|||   |||
Db     2155 GGCTTACAGCCAGCATTCTCCGTAATATTTCAAAAGTTGTAGGGGTTTTTACAAATGGT 2214

Qy      125 AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeu 144
                           |||   :::   ::::::   ::::::|||::::::|||:::||||||
Db     2215 GGCCCCCCTGAAACCCAGCGGCAGAACAGCAACTTGTTTTCCTTGTTGTTTCTGATCCTT 2274

Qy      145 GlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGlu 164
            |||:::|||   |||::::::::|||:::||||||          :::|||:::|||||||||
Db     2275 GGGATCATTTCTTTCATTACATTTTTTCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAG 2334

Qy      165 IleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAla 184
            |||||||||    |||||||||::::::      ||||||:::||||||    ||||||::::::
Db     2335 ATCCTCACCAAGCGACTCCGATACATGGTTTTCAAATCCATGCTGAGACAGGATGTGAGC 2394

Qy      185 TrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIle 204
            |||||||||:::   :::|||:::||||||   |||||||||   ||||||   |||
Db     2395 TGGTTTGATGACCCTAAAAACACCACCGGAGCACTGACCACCAGGCTCGCCAACGATGCT 2454
```

```
Qy    205 AlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsn 224
Db   2455 GCTCAAGTGAAAGGGGCTACAGGGTCTAGGCTTGCTGTGATTTTCCAGAACATAGCAAAT 2514

Qy    225 MetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeu 244
Db   2515 CTTGGGACAGGAATCATCATATCCCTAATCTATGGCTGGCAACTAACACTTTTACTCTTA 2574

Qy    245 SerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPhe 264
Db   2575 GCAATTGTACCCATCATTGCGATAGCTGGAGTGGTTGAAATGAAAATGTTGTCTGGACAA 2634

Qy    265 AlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGlu 284
Db   2635 GCACTGAAAGATAAGAAGGAACTAGAAGGTTCTGGAAAGATTGCTACGGAAGCAATTGAA 2694

Qy    285 AsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGlu 304
Db   2695 AACTTCCGCACTGTTGTCTCTTTGACTCGGGAGCAGAAGTTTGAAACCATGTATGCCCAG 2754

Qy    305 MetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyr 324
Db   2755 AGCTTGCAGATACCATACAGAAATGCGATGAAGAAAGCACACGTGTTTGGGATCACGTTC 2814

Qy    325 AlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyr 344
Db   2815 TCCTTCACCCAGGCCATGATGTATTTTTCTTATGCTGCTTGTTTCCGGTTCGGTGCCTAC 2874

Qy    345 LeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAla 364
Db   2875 TTGGTGACACAACAACTCATGACTTTTGAAAATGTTCTGTTAGTATTCTCAGCTATTGTC 2934

Qy    365 TyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLys 384
Db   2935 TTTGGTGCCATGGCAGTGGGGCAGGTCAGTTCATTCGCTCCTGACTATGCGAAAGCAACA 2994

Qy    385 SerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSer 404
Db   2995 GTGTCAGCATCCCACATCATCAGGATCATTGAGAAAACCCCCGAGATTGACAGCTACAGC 3054

Qy    405 GlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhe 424
Db   3055 ACGCAAGGCCTAAAGCCGAATATGTTGGAAGGAAATGTGCAATTTAGTGGAGTCGTGTTC 3114

Qy    425 PheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArg 444
Db   3115 AACTATCCCACCCGACCCAGCATCCCAGTGCTTCAGGGGCTGAGCCTTGAGGTGAAGAAG 3174

Qy    445 GlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeu 464
Db   3175 GGCCAGACGCTGGCCCTGGTGGGCAGCAGTGGCTGCGGGAAGAGCACAGTGGTCCAGCTG 3234

Qy    465 LeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLys 484
Db   3235 CTCGAGCGCTTCTACGACCCCATGGCTGGATCAGTGTTTCTAGATGGCAAAGAAATAAAG 3294

Qy    485 GluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeu 504
Db   3295 CAACTGAATGTCCAGTGGCTCCGAGCACAGCTGGGCATTGTGTCCCAAGAGCCCATTCTC 3354

Qy    505 PheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeu 524
Db   3355 TTTGACTGCAGCATCGCAGAGAACATTGCCTACGGAGACAACAGCCGGGTCGTGTCTTAT 3414

Qy    525 AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuPro 544
Db   3415 GAGGAGATTGTGAGGGCAGCCAAGGAGGCCAACATCCACCAGTTCATCGACTCGCTACCT 3474

Qy    545 GluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGln 564
Db   3475 GATAAATACAACACCAGAGTAGGAGACAAAGGCACTCAGCTGTCGGGTGGGCAGAAGCAG 3534

Qy    565 ArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAla 584
Db   3535 CGCATCGCCATCGCACGCGCCCTCGTCAGACAGCCTCACATTTTACTTCTGGACGAAGCA 3594

Qy    585 ThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArg 604
Db   3595 ACATCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAGGAAGCGCTGGACAAAGCCAGG 3654

Qy    605 ThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeu 624
Db   3655 GAAGGCCGCACCTGCATTGTGATCGCTCACCGCCTGTCCACCATCCAGAACGCGGACTTG 3714

Qy    625 IleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArg 644
Db   3715 ATCGTGGTGATTCAGAACGGCAAGGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCG 3774

Qy    645 AsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db   3775 CAGAAGGGCATCTACTTCTCAATGGTCAGTGTGCAGGCT 3813
```

Search completed: March 31, 2003, 15:57:49
Job time : 327.871 secs

```
OM protein - nucleic search, using frame_plus_p2n model

Run on:          March 30, 2003, 01:24:25 ; Search time 8908.97 Seconds
                                            (without alignments)
                                            2152.747 Million cell updates/sec

Title:           US-09-873-409-1
Perfect score:   3323
Sequence:        1 MLAEKGAHAELMAKRGLYYS..........QELLRNRDIYFKLVNAQSVQ 659

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters:       4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115418_19228/app_query.fasta_1.7544
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_54887_@runat_27032003_115418_19228 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       GenEmbl:*
                 1:  gb_ba:*
                 2:  gb_htg:*
                 3:  gb_in:*
                 4:  gb_om:*
                 5:  gb_ov:*
                 6:  gb_pat:*
                 7:  gb_ph:*
                 8:  gb_pl:*
                 9:  gb_pr:*
                 10:  gb_ro:*
                 11:  gb_sts:*
                 12:  gb_sy:*
                 13:  gb_un:*
                 14:  gb_vi:*
                 15:  em_ba:*
                 16:  em_fun:*
                 17:  em_hum:*
                 18:  em_in:*
                 19:  em_mu:*
                 20:  em_om:*
                 21:  em_or:*
                 22:  em_ov:*
                 23:  em_pat:*
                 24:  em_ph:*
                 25:  em_pl:*
                 26:  em_ro:*
                 27:  em_sts:*
                 28:  em_un:*
                 29:  em_vi:*
                 30:  em_htg_hum:*
                 31:  em_htg_inv:*
                 32:  em_htg_other:*
                 33:  em_htg_mus:*
                 34:  em_htg_pln:*
                 35:  em_htg_rod:*
                 36:  em_htg_mam:*
                 37:  em_htg_vrt:*
                 38:  em_sy:*
                 39:  em_htgo_hum:*
                 40:  em_htgo_mus:*
                 41:  em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3323	100.0	2066	6	AX339027	AX339027 Sequence
2	3323	100.0	2856	6	AX339028	AX339028 Sequence
3	3323	100.0	3177	6	AX339030	AX339030 Sequence
4	3323	100.0	3621	6	AX339032	AX339032 Sequence
5	3323	100.0	3702	6	AX339031	AX339031 Sequence
6	3163	95.2	3699	6	AX478104	AX478104 Sequence
7	1924.5	57.9	3924	6	AX024455	AX024455 Sequence
8	1924.5	57.9	3924	6	AX409652	AX409652 Sequence
9	1924.5	57.9	3924	9	HUMMDR3	M23234 Human membr
10	1916	57.7	3987	10	CRUPGP1185	M59254 Chinese ham
11	1916	57.7	4296	10	CRUPGPI	M60040 C.griseus P
12	1916	57.7	4304	10	CRUPGP1165	M59253 Chinese ham
13	1915.5	57.6	4084	10	MUSMDRA	J03398 Mouse mdr g
14	1911	57.5	4018	5	GGA9799	AJ009799 Gallus ga
15	1893	57.0	3858	4	OAU78609	U78609 Ovis aries
16	1892	56.9	4927	10	AF257746	AF257746 Rattus no
17	1884	56.7	2752	5	AF099732	AF099732 Fundulus
18	1882.5	56.7	3912	6	AX401884	AX401884 Sequence
19	1882.5	56.7	3912	10	RATPGLYCO	L15079 Rat p-glyco
20	1879	56.5	4323	10	AF286167	AF286167 Rattus no
21	1876.5	56.5	3905	10	CRUPGPIII	M60042 C.griseus P
22	1873	56.4	4186	6	AX108654	AX108654 Sequence
23	1873	56.4	4195	6	AX108656	AX108656 Sequence
24	1872	56.3	2726	6	I33621	I33621 Sequence 1
25	1872	56.3	3840	6	AX481416	AX481416 Sequence
26	1872	56.3	3860	6	AX322787	AX322787 Sequence
27	1872	56.3	3860	6	AX322789	AX322789 Sequence
28	1872	56.3	3988	6	AX024454	AX024454 Sequence
29	1872	56.3	4192	9	AF016535	AF016535 Homo sapi
30	1872	56.3	4264	6	AR051647	AR051647 Sequence
31	1872	56.3	4264	6	AR051650	AR051650 Sequence
32	1872	56.3	4646	6	AX336420	AX336420 Sequence
33	1872	56.3	4646	6	AX336708	AX336708 Sequence
34	1872	56.3	4646	6	AX391099	AX391099 Sequence
35	1872	56.3	4646	6	I49610	I49610 Sequence 2
36	1872	56.3	4646	9	HUMMDR1	M14758 Homo sapien
37	1872	56.3	4669	6	I08557	I08557 Sequence 3
38	1872	56.3	6505	6	AR028671	AR028671 Sequence
39	1872	56.3	8630	6	AX012320	AX012320 Sequence
40	1872	56.3	8630	6	AX012321	AX012321 Sequence
41	1872	56.3	9318	6	AR028672	AR028672 Sequence
42	1871	56.3	3682	10	MUSPGLY1A	M24417 Mouse phosp
43	1871	56.3	4788	6	AX322793	AX322793 Sequence
44	1871	56.3	4924	10	MUSMDR1A	M33581 Mouse P-gly
45	1870	56.3	4279	6	AX105057	AX105057 Sequence

ALIGNMENTS

RESULT 1

```
AX339027
LOCUS       AX339027                2066 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 9 from Patent WO0194400.
ACCESSION   AX339027
VERSION     AX339027.1  GI:18129119
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 9 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..2066
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      647 a    403 c    470 g    546 t
ORIGIN

Alignment Scores:
Pred. No.:              1.8e-259     Length:        2066
Score:                  3323.00      Matches:       659
Percent Similarity:     100.00%      Conservative:  0
Best Local Similarity:  100.00%      Mismatches:    0
Query Match:            100.00%      Indels:        0
DB:                     6            Gaps:          0

US-09-873-409-1 (1-659) x AX339027 (1-2066)

Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     87 ATGCTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCA 146

Qy     21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    147 CTTGTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCT 206

Qy     41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    207 ACTGAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTC 266

Qy     61 IleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    267 ATTGACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTA 326

Qy     81 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    327 TTAAAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCT 386

Qy    101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    387 TCTGTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACC 446

Qy    121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    447 ATGTTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATA 506

Qy    141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    507 TTCGTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGC 566

Qy    161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    567 AGAGCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATAT 626

Qy    181 GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    627 CAGGATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTA 686

Qy    201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    687 GCCATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAA 746

Qy    221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    747 AATGCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACA 806

Qy    241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    807 TTCCTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCA 866

Qy    261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    867 ATGACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACT 926

Qy    281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    927 GAAGCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAA 986

Qy    301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    987 ATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATT 1046

Qy    321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1047 GGAAGCTGTTATGCATTCAGCCATGCCTTTATATATTTTGCCTATGCAGCAGGGTTTCGA 1106

Qy    341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1107 TTTGGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTT 1166

Qy    361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1167 ACTGCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATAT 1226

Qy    381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1227 TCCAAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATA 1286

Qy    401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1287 GACAGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGA 1346

Qy    421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1347 GAAGTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTC 1406

Qy    441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1407 AGTATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGAAAAAGCACT 1466

Qy    461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1467 TCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGT 1526

Qy    481 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1527 GTGGATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAA 1586

Qy    501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1587 GAGCCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGT 1646

Qy    521 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1647 GTGGTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATT 1706

Qy    541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1707 GAAGGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGC 1766

Qy    561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580
```

```
               |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1767 GGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTG 1826

Qy       581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1827 TTGGATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTT 1886

Qy       601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1887 GATAAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAG 1946

Qy       621 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1947 AACGCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAA 2006

Qy       641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2007 GAGCTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 2063

RESULT 2
AX339028
LOCUS       AX339028                2856 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 10 from Patent WO0194400.
ACCESSION   AX339028
VERSION     AX339028.1  GI:18129120
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 10 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..2856
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      888 a    540 c    652 g    776 t
ORIGIN

Alignment Scores:
Pred. No.:              2.77e-259      Length:        2856
Score:                  3323.00        Matches:       659
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     6              Gaps:          0

US-09-873-409-1 (1-659) x AX339028 (1-2856)

Qy         1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       877 ATGCTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCA 936

Qy        21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       937 CTTGTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCT 996

Qy        41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       997 ACTGAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTC 1056

Qy        61 IleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1057 ATTGACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTA 1116

Qy        81 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1117 TTAAAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCT 1176

Qy       101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1177 TCTGTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACC 1236

Qy       121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1237 ATGTTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATA 1296

Qy       141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1297 TTCGTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGC 1356

Qy       161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1357 AGAGCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATAT 1416

Qy       181 GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1417 CAGGATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTA 1476

Qy       201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1477 GCCATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAA 1536

Qy       221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1537 AATGCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACA 1596

Qy       241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1597 TTCCTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCA 1656

Qy       261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1657 ATGACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACT 1716

Qy       281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1717 GAAGCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAA 1776

Qy       301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1777 ATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATT 1836

Qy       321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1837 GGAAGCTGTTATGCATTCAGCCATGCCTTTATATATTTGCCTATGCAGCAGGGTTTCGA 1896

Qy       341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1897 TTTGGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTT 1956

Qy       361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1957 ACTGCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATAT 2016

Qy       381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2017 TCCAAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATA 2076

Qy       401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2077 GACAGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGA 2136

Qy       421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2137 GAAGTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTC 2196

Qy       441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2197 AGTATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACT 2256

Qy       461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2257 TCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGT 2316

Qy      481 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2317 GTGGATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAA 2376

Qy      501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2377 GAGCCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGT 2436

Qy      521 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2437 GTGGTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATT 2496

Qy      541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2497 GAAGGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGC 2556

Qy      561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2557 GGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTG 2616

Qy      581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2617 TTGGATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTT 2676

Qy      601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2677 GATAAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAG 2736

Qy      621 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2737 AACGCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAA 2796

Qy      641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2797 GAGCTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 2853

RESULT 3
AX339030
LOCUS       AX339030                3177 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 12 from Patent WO0194400.
ACCESSION   AX339030
VERSION     AX339030.1  GI:18129122
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 12 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..3177
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      970 a    601 c    735 g    870 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              3.2e-259       Length:         3177
Score:                  3323.00        Matches:        659
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            100.00%        Indels:         0
DB:                     6              Gaps:           0

US-09-873-409-1 (1-659) x AX339030 (1-3177)

Qy        1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1198 ATGCTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCA 1257

Qy       21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1258 CTTGTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCT 1317

Qy       41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1318 ACTGAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTC 1377

Qy       61 IleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1378 ATTGACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTA 1437

Qy       81 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1438 TTAAAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCT 1497

Qy      101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1498 TCTGTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACC 1557

Qy      121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1558 ATGTTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATA 1617

Qy      141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1618 TTCGTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGC 1677

Qy      161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1678 AGAGCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATAT 1737

Qy      181 GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1738 CAGGATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTA 1797

Qy      201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1798 GCCATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAA 1857

Qy      221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1858 AATGCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACA 1917

Qy      241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1918 TTCCTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCA 1977

Qy      261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1978 ATGACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACT 2037

Qy      281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2038 GAAGCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAA 2097

Qy      301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2098 ATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATT 2157

Qy      321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2158 GGAAGCTGTTATGCATTCAGCCATGCCTTTATATATTTGCCTATGCAGCAGGGTTTCGA 2217

Qy      341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2218 TTTGGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTT 2277

Qy      361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2278 ACTGCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATAT 2337

Qy       381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2338 TCCAAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATA 2397

Qy       401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2398 GACAGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGA 2457

Qy       421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2458 GAAGTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTC 2517

Qy       441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2518 AGTATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACT 2577

Qy       461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2578 TCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGT 2637

Qy       481 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2638 GTGGATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAA 2697

Qy       501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2698 GAGCCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGT 2757

Qy       521 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2758 GTGGTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATT 2817

Qy       541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2818 GAAGGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGC 2877

Qy       561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2878 GGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTG 2937

Qy       581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2938 TTGGATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTT 2997

Qy       601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2998 GATAAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAG 3057

Qy       621 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3058 AACGCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAA 3117

Qy       641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3118 GAGCTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3174
```

```
RESULT 4
AX339032
LOCUS       AX339032                3621 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 14 from Patent WO0194400.
ACCESSION   AX339032
VERSION     AX339032.1  GI:18129124
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 14 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..3621
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1081 a    696 c    842 g   1002 t
ORIGIN

Alignment Scores:
Pred. No.:              3.81e-259       Length:         3621
Score:                  3323.00         Matches:        659
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            100.00%         Indels:         0
DB:                     6               Gaps:           0

US-09-873-409-1 (1-659) x AX339032 (1-3621)
```

```
Qy         1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1642 ATGCTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCA 1701

Qy        21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1702 CTTGTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCT 1761

Qy        41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1762 ACTGAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTC 1821

Qy        61 IleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1822 ATTGACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTA 1881

Qy        81 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1882 TTAAAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCT 1941

Qy       101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1942 TCTGTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACC 2001

Qy       121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2002 ATGTTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATA 2061

Qy       141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2062 TTCGTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGC 2121

Qy       161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2122 AGAGCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATAT 2181

Qy       181 GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2182 CAGGATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTA 2241

Qy       201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2242 GCCATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAA 2301

Qy       221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2302 AATGCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACA 2361

Qy       241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2362 TTCCTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCA 2421

Qy       261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2422 ATGACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACT 2481

Qy      281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2482 GAAGCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAA 2541

Qy      301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2542 ATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATT 2601

Qy      321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2602 GGAAGCTGTTATGCATTCAGCCATGCCTTTATATATTTTGCCTATGCAGCAGGGTTTCGA 2661

Qy      341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2662 TTTGGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTT 2721

Qy      361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2722 ACTGCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATAT 2781

Qy      381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2782 TCCAAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATA 2841

Qy      401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2842 GACAGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGA 2901

Qy      421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2902 GAAGTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTC 2961

Qy      441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2962 AGTATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACT 3021

Qy      461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3022 TCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGT 3081

Qy      481 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3082 GTGGATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAA 3141

Qy      501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3142 GAGCCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGT 3201

Qy      521 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3202 GTGGTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATT 3261

Qy      541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3262 GAAGGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGC 3321

Qy      561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3322 GGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTG 3381

Qy      581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3382 TTGGATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTT 3441

Qy      601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3442 GATAAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAG 3501

Qy      621 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3502 AACGCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAA 3561

Qy      641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3562 GAGCTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3618
```

```
RESULT 5
AX339031
LOCUS       AX339031                3702 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 13 from Patent WO0194400.
ACCESSION   AX339031
VERSION     AX339031.1  GI:18129123
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 13 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..3702
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1098 a    708 c    856 g   1039 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              3.92e-259      Length:          3702
Score:                  3323.00        Matches:         659
Percent Similarity:     100.00%        Conservative:    0
Best Local Similarity:  100.00%        Mismatches:      0
Query Match:            100.00%        Indels:          0
DB:                     6              Gaps:            0

US-09-873-409-1 (1-659) x AX339031 (1-3702)

Qy        1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1723 ATGCTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCA 1782

Qy       21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1783 CTTGTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCT 1842

Qy       41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1843 ACTGAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTC 1902

Qy       61 IleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1903 ATTGACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTA 1962

Qy       81 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1963 TTAAAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCT 2022

Qy      101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2023 TCTGTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACC 2082

Qy      121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2083 ATGTTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATA 2142

Qy      141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2143 TTCGTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGC 2202

Qy      161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180
```

```
                 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2203 AGAGCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATAT 2262

Qy       181 GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2263 CAGGATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTA 2322

Qy       201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2323 GCCATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAA 2382

Qy       221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2383 AATGCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACA 2442

Qy       241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2443 TTCCTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCA 2502

Qy       261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2503 ATGACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACT 2562

Qy       281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2563 GAAGCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAA 2622

Qy       301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2623 ATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATT 2682

Qy       321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2683 GGAAGCTGTTATGCATTCAGCCATGCCTTTATATATTTGCCTATGCAGCAGGGTTTCGA 2742

Qy       341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2743 TTTGGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTT 2802

Qy       361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2803 ACTGCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATAT 2862

Qy       381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2863 TCCAAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATA 2922

Qy       401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2923 GACAGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGA 2982

Qy       421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      2983 GAAGTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTC 3042

Qy       441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3043 AGTATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACT 3102

Qy       461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3103 TCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGT 3162

Qy       481 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3163 GTGGATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAA 3222

Qy       501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3223 GAGCCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGT 3282

Qy       521 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3283 GTGGTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATT 3342

Qy       541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3343 GAAGGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGC 3402

Qy       561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3403 GGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTG 3462

Qy       581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3463 TTGGATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTT 3522

Qy       601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3523 GATAAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAG 3582

Qy       621 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 640
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3583 AACGCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAA 3642

Qy       641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3643 GAGCTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3699

RESULT 6
AX478104
LOCUS       AX478104                3699 bp    DNA     linear   PAT 12-AUG-2002
DEFINITION  Sequence 34 from Patent WO0240541.
ACCESSION   AX478104
VERSION     AX478104.1  GI:22217064
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Tang,Y.T., Yue,H., Nguyen,D.B., Hafalia,A.J., Elliott,V.S., Lu,Y.,
            Walia,N.K., Yao,M.G., Baughn,M.R., Gandhi,A.R., Ding,L.,
            Sanjanwala,M., Ramkumar,J., Arvizu,C., Gietzen,K.J., Lal,P.G.,
            Azimzai,Y., Khan,F.A., Thangavelu,K., Thornton,M., Lu,D.A.,
            Tribouley,C.M., Warren,B.A., Ison,C.H., Das,D., Raumann,B.E.,
            Policky,J.L. and Kearney,L.
  TITLE     Transporters and ion channels
  JOURNAL   Patent: WO 0240541-A 34 23-MAY-2002;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..3699
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 7472030CB1"
BASE COUNT     1116 a    707 c    860 g   1016 t
ORIGIN

Alignment Scores:
Pred. No.:              3.44e-246       Length:        3699
Score:                  3163.00         Matches:       632
Percent Similarity:     96.05%          Conservative:  1
Best Local Similarity:  95.90%          Mismatches:    0
Query Match:            95.19%          Indels:        26
DB:                     6               Gaps:          1

US-09-873-409-1 (1-659) x AX478104 (1-3699)

Qy         1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1798 ATGCTGGCGGAGAAAGGAGCGCATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCA 1857

Qy        21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1858 CTTGTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCT 1917
```

```
Qy      41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1918 ACTGAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTC 1977

Qy      61 IleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1978 ATTGACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTA 2037

Qy      81 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2038 TTAAAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCT 2097

Qy     101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2098 TCTGTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACC 2157

Qy     121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2158 ATGTTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATA 2217

Qy     141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160
           ||||||||||||||||||||||||||||||||||||||||||
Db    2218 TTCGTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATG------------------ 2259

Qy     161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180

Db    2259 ------------------------------------------------------------ 2259

Qy     181 GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2260 CAGGATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTA 2319

Qy     201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2320 GCCATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAA 2379

Qy     221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2380 AATGCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACA 2439

Qy     241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2440 TTCCTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCA 2499

Qy     261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2500 ATGACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACT 2559

Qy     281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2560 GAAGCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAA 2619

Qy     301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2620 ATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATT 2679

Qy     321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2680 GGAAGCTGTTATGCATTCAGCCATGCCTTTATATATTTTGCCTATGCGGCAGGGTTTCGA 2739

Qy     341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2740 TTTGGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTT 2799

Qy     361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380
           ||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db    2800 ACTGCAATTGCATATGGAGCTATGGCCATCGGAGAAACGCTCGTTTTGGCTCCTGAATAT 2859

Qy     381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2860 TCCAAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATA 2919

Qy     401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2920 GACAGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGA 2979

Qy     421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2980 GAAGTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTC 3039

Qy     441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3040 AGTATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGAAAAAGCACT 3099

Qy     461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3100 TCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGT 3159

Qy     481 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3160 GTGGATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAA 3219

Qy     501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3220 GAGCCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGT 3279

Qy     521 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3280 GTGGTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATT 3339

Qy     541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3340 GAAGGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGC 3399

Qy     561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3400 GGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTG 3459

Qy     581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3460 TTGGATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTT 3519

Qy     601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3520 GATAAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAG 3579

Qy     621 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3580 AACGCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAA 3639

Qy     641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3640 GAGCTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3696

RESULT 7
AX024455
LOCUS       AX024455                3924 bp    DNA     linear   PAT 15-SEP-2000
DEFINITION  Sequence 2 from Patent DE19857750.
ACCESSION   AX024455
VERSION     AX024455.1  GI:10184623
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3924)
  AUTHORS   Stremmel,W.
  JOURNAL   Patent: DE 19857750-A 2 24-FEB-2000;
            STREMMEL WOLFGANG (DE)
FEATURES             Location/Qualifiers
     source          1..3924
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="MDR3/menschliche MDR-Genfamilie"
BASE COUNT     1145 a    790 c    977 g   1012 t
ORIGIN
```

```
Alignment Scores:
Pred. No.:              5.8e-146      Length:        3924
Score:                  1924.50       Matches:       372
Percent Similarity:     74.10%        Conservative:  123
Best Local Similarity:  55.69%        Mismatches:    162
Query Match:            57.91%        Indels:        11
DB:                     6             Gaps:          3

US-09-873-409-1 (1-659) x AX024455 (1-3924)

Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db   1851 GTAATTGTGGAGCAAGGAAGCCACAGCGAACTGATGAAGAAGGAAGGGGTGTACTTCAAA 1910

Qy     21 LeuVal---------------------------MetSerGlnAspIleLysLysAlaAsp 31
Db   1911 CTTGTCAACATGCAGACATCAGGAAGCCAGATCCAGTCAGAAGAATTTGAACTAAATGAT 1970

Qy     32 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 51
Db   1971 GAAAAGGCTGCCACTAGAATGGCCCCAAATGGCTGGAAATCTCGCCTATTTAGGCATTCT 2030

Qy     52 SerValLysSerIleLys---SerAspPheIleAspLysAla---GluGluSerThrGln 69
Db   2031 ACTCAGAAAAACCTTAAAAATTCACAAATGTGTCAGAAGAGCCTTGATGTGGAAACCGAT 2090

Qy     70 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 89
Db   2091 GGACTTGAAGCAAATGTGCCACCAGTGTCCTTTCTGAAGGTCCTGAAACTGAATAAAACA 2150

Qy     90 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 109
Db   2151 GAATGGCCCTACTTTGTCGTGGGAACAGTATGTGCCATTGCCAATGGGGGGCTTCAGCCG 2210

Qy    110 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 129
Db   2211 GCATTTTCAGTCATATTCTCAGAGATCATAGCGATTTTTGGACCAGGCGATGATGCAGTG 2270

Qy    130 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 149
Db   2271 AAGCAGCAGAAGTGCAACATATTCTCTTTGATTTTCTTATTTCTGGGAATTATTTCTTTT 2330

Qy    150 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 169
Db   2331 TTTACTTTCTTCCTTCAGGGTTTCACGTTTGGGAAAGCTGGCGAGATCCTCACCAGAAGA 2390

Qy    170 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLys 189
Db   2391 CTGCGGTCAATGGCTTTTAAAGCAATGCTAAGACAGGACATGAGCTGGTTTGATGACCAT 2450

Qy    190 GluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 209
Db   2451 AAAAACAGTACTGGTGCACTTTCTACAAGACTTGCCACAGATGCTGCCCAAGTCCAAGGA 2510

Qy    210 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 229
Db   2511 GCCACAGGAACCAGGTTGGCTTTAATTGCACAGAATATAGCTAACCTTGGAACTGGTATT 2570

Qy    230 IleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVal 249
Db   2571 ATCATATCATTTATCTACGGTTGGCAGTTAACCCTATTGCTATTAGCAGTTGTTCCAATT 2630

Qy    250 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 269
Db   2631 ATTGCTGTGTCAGGAATTGTTGAAATGAAATTGTTGGCTGGAAATGCCAAAAGAGATAAA 2690

Qy    270 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIle 289
Db   2691 AAAGAACTGGAAGCTGCTGGAAAGATTGCAACAGAGGCAATAGAAAATATTAGGACAGTT 2750

Qy    290 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGln 309
Db   2751 GTGTCTTTGACCCAGGAAAGAAAATTTGAATCAATGTATGTTGAAAAATTGTATGGACCT 2810

Qy    310 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla 329
Db   2811 TACAGGAATTCTGTGCAGAAGGCACACATCTATGGAATTACTTTTAGTATCTCACAAGCA 2870

Qy    330 PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGly 349
Db   2871 TTTATGTATTTTTCCTATGCCGGTTGTTTTCGATTTGGTGCATATCTCATTGTGAATGGA 2930

Qy    350 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAla 369
Db   2931 CATATGCGCTTCAGAGATGTTATTCTGGTGTTTTCTGCAATTGTATTTGGTGCAGTGGCT 2990

Qy    370 IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHis 389
Db   2991 CTAGGACATGCCAGTTCATTTGCTCCAGACTATGCTAAAGCTAAGCTGTCTGCAGCCCAC 3050

Qy    390 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLys 409
Db   3051 TTATTCATGCTGTTTGAAAGACAACCTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAG 3110

Qy    410 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 429
Db   3111 CCTGATAAATTTGAAGGAAATATAACATTTAATGAAGTCGTGTTCAACTATCCCACCCGA 3170

Qy    430 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 449
Db   3171 GCAAACGTGCCAGTGCTTCAGGGGCTGAGCCTGGAGGTGAAGAAAGGCCAGACACTAGCC 3230

Qy    450 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr 469
Db   3231 CTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACGGTGGTCCAGCTCCTGGAGCGGTTCTAC 3290

Qy    470 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 489
Db   3291 GACCCCTTGGCGGGGACAGTGCTTCTCGATGGTCAAGAAGCAAAGAAACTCAATGTCCAG 3350

Qy    490 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 509
Db   3351 TGGCTCAGAGCTCAACTCGGAATCGTGTCTCAGGAGCCTATCCTATTTGACTGCAGCATT 3410

Qy    510 AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu 529
Db   3411 GCCGAGAATATTGCCTATGGAGACAACAGCCGGGTTGTATCACAGGATGAAATTGTGAGT 3470

Qy    530 AlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThr 549
Db   3471 GCAGCCAAAGCTGCCAACATACATCCTTTCATCGAGACGTTACCCCACAAATATGAAACA 3530

Qy    550 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAla 569
Db   3531 AGAGTGGGAGATAAGGGGACTCAGCTCTCAGGAGGTCAAAAACAGAGGATTGCTATTGCC 3590

Qy    570 ArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsp 589
Db   3591 CGAGCCCTCATCAGACAACCTCAAATCCTCCTGTTGGATGAAGCTACATCAGCTCTGGAT 3650

Qy    590 AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys 609
Db   3651 ACTGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGC 3710

Qy    610 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 629
Db   3711 ATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAG 3770

Qy    630 AsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyr 649
Db   3771 AATGGGAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTAT 3830

Qy    650 PheLysLeuValAsnAlaGlnSer 657
Db   3831 TTTTCAATGGTCAGTGTCCAGGCT 3854
```

```
RESULT 8
AX409652
LOCUS       AX409652                3924 bp    DNA     linear   PAT 14-JUN-2002
DEFINITION  Sequence 2299 from Patent WO0229103.
ACCESSION   AX409652
VERSION     AX409652.1  GI:21442357
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  .
  AUTHORS   Alvares,C., Horne,D., Peres-da-Silva,S. and Vockley,J.G.
  TITLE     Gene expression profiles in liver cancer
  JOURNAL   Patent: WO 0229103-A 2299 11-APR-2002;
            GENE LOGIC INC (US)
FEATURES             Location/Qualifiers
     source          1. .3924
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="EMBL/GenBank Accession No. M23234"
BASE COUNT     1145 a    790 c    977 g   1012 t
ORIGIN

Alignment Scores:
Pred. No.:              5.8e-146      Length:        3924
Score:                  1924.50       Matches:       372
Percent Similarity:     74.10%        Conservative:  123
Best Local Similarity:  55.69%        Mismatches:    162
Query Match:            57.91%        Indels:        11
DB:                     6             Gaps:          3

US-09-873-409-1 (1-659) x AX409652 (1-3924)

Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db   1851 GTAATTGTGGAGCAAGGAAGCCACAGCGAACTGATGAAGAAGGAAGGGGTGTACTTCAAA 1910

Qy     21 LeuVal---------------------------MetSerGlnAspIleLysLysAlaAsp 31
Db   1911 CTTGTCAACATGCAGACATCAGGAAGCCAGATCCAGTCAGAAGAATTTGAACTAAATGAT 1970

Qy     32 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 51
Db   1971 GAAAAGGCTGCCACTAGAATGGCCCCAAATGGCTGGAAATCTCGCCTATTTAGGCATTCT 2030

Qy     52 SerValLysSerIleLys---SerAspPheIleAspLysAla---GluGluSerThrGln 69
Db   2031 ACTCAGAAAAACCTTAAAAATTCACAAATGTGTCAGAAGAGCCTTGATGTGGAAACCGAT 2090

Qy     70 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 89
Db   2091 GGACTTGAAGCAAATGTGCCACCAGTGTCCTTTCTGAAGGTCCTGAAACTGAATAAAACA 2150

Qy     90 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 109
Db   2151 GAATGGCCCTACTTTGTCGTGGGAACAGTATGTGCCATTGCCAATGGGGGGCTTCAGCCG 2210

Qy    110 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 129
Db   2211 GCATTTTCAGTCATATTCTCAGAGATCATAGCGATTTTTGGACCAGGCGATGATGCAGTG 2270

Qy    130 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 149
Db   2271 AAGCAGCAGAAGTGCAACATATTCTCTTTGATTTTCTTATTTCTGGGAATTATTTCTTTT 2330

Qy    150 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 169
Db   2331 TTTACTTTCTTCCTTCAGGGTTTCACGTTTGGGAAAGCTGGCGAGATCCTCACCAGAAGA 2390

Qy    170 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLys 189
Db   2391 CTGCGGTCAATGGCTTTTAAAGCAATGCTAAGACAGGACATGAGCTGGTTTGATGACCAT 2450

Qy    190 GluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 209
Db   2451 AAAAACAGTACTGGTGCACTTTCTACAAGACTTGCCACAGATGCTGCCCAAGTCCAAGGA 2510

Qy    210 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 229
Db   2511 GCCACAGGAACCAGGTTGGCTTTAATTGCACAGAATATAGCTAACCTTGGAACTGGTATT 2570

Qy    230 IleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVal 249
Db   2571 ATCATATCATTTATCTACGGTTGGCAGTTAACCCTATTGCTATTAGCAGTTGTTCCAATT 2630

Qy    250 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 269
Db   2631 ATTGCTGTGTCAGGAATTGTTGAAATGAAATTGTTGGCTGGAAATGCCAAAAGAGATAAA 2690

Qy    270 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIle 289
Db   2691 AAAGAACTGGAAGCTGCTGGAAAGATTGCAACAGAGGCAATAGAAAATATTAGGACAGTT 2750

Qy    290 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGln 309
Db   2751 GTGTCTTTGACCCAGGAAAGAAAATTTGAATCAATGTATGTTGAAAAATTGTATGGACCT 2810

Qy    310 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla 329
Db   2811 TACAGGAATTCTGTGCAGAAGGCACACATCTATGGAATTACTTTTAGTATCTCACAAGCA 2870

Qy    330 PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGly 349
Db   2871 TTTATGTATTTTTCCTATGCCGGTTGTTTTCGATTTGGTGCATATCTCATTGTGAATGGA 2930

Qy    350 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAla 369
Db   2931 CATATGCGCTTCAGAGATGTTATTCTGGTGTTTTCTGCAATTGTATTTGGTGCAGTGGCT 2990

Qy    370 IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHis 389
Db   2991 CTAGGACATGCCAGTTCATTTGCTCCAGACTATGCTAAAGCTAAGCTGTCTGCAGCCCAC 3050

Qy    390 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLys 409
Db   3051 TTATTCATGCTGTTTGAAAGACAACCTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAG 3110

Qy    410 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 429
Db   3111 CCTGATAAATTTGAAGGAAATATAACATTTAATGAAGTCGTGTTCAACTATCCCACCCGA 3170

Qy    430 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 449
Db   3171 GCAAACGTGCCAGTGCTTCAGGGGCTGAGCCTGGAGGTGAAGAAGGCCAGACACTAGCC 3230

Qy    450 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr 469
Db   3231 CTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACGGTGGTCCAGCTCCTGGAGCGGTTCTAC 3290

Qy    470 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 489
Db   3291 GACCCCTTGGCGGGGACAGTGCTTCTCGATGGTCAAGAAGCAAAGAAACTCAATGTCCAG 3350

Qy    490 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 509
Db   3351 TGGCTCAGAGCTCAACTCGGAATCGTGTCTCAGGAGCCTATCCTATTTGACTGCAGCATT 3410

Qy    510 AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu 529
Db   3411 GCCGAGAATATTGCCTATGGAGACAACAGCCGGGTTGTATCACAGGATGAAATTGTGAGT 3470

Qy    530 AlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThr 549
Db   3471 GCAGCCAAAGCTGCCAACATACATCCTTTCATCGAGACGTTACCCCACAAATATGAAACA 3530
```

```
Qy     550 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAla 569
           :::|||||   |||||   |||||||||||||||||||||||||:::||||||||
Db    3531 AGAGTGGGAGATAAGGGGACTCAGCTCTCAGGAGGTCAAAAACAGAGGATTGCTATTGCC 3590

Qy     570 ArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsp 589
           ||||||||:::::::|||:::|||||||||||||||||||||||||||||||||||
Db    3591 CGAGCCCTCATCAGACAACCTCAAATCCTCCTGTTGGATGAAGCTACATCAGCTCTGGAT 3650

Qy     590 AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys 609
              :::||||||||||||||||   ||||||||||||||||   |||||||||||
Db    3651 ACTGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGC 3710

Qy     610 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 629
           :::|||:::   ||||||||||   ||||||||||||||||||||||||||
Db    3711 ATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAG 3770

Qy     630 AsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyr 649
           ||||||::::::||||||   ||||||||||:::||||||      :::   ||||||
Db    3771 AATGGGAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTAT 3830

Qy     650 PheLysLeuValAsnAlaGlnSer 657
           |||   :::|||:::   |||:::
Db    3831 TTTCAATGGTCAGTGTCCAGGCT 3854
```

```
RESULT 9
HUMMDR3
LOCUS       HUMMDR3                  3924 bp    mRNA    linear   PRI 11-JUN-1993
DEFINITION  Human membrane glycoprotein P (mdr3) mRNA, complete cds.
ACCESSION   M23234
VERSION     M23234.1  GI:187501
KEYWORDS    P-glycoprotein; membrane glycoprotein.
SOURCE      Human liver, cDNA to mRNA, clone 3.27.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 3924)
  AUTHORS   van der Bliek,A.M., Kooiman,P.M., Schneider,C. and Borst,P.
  TITLE     Sequence of mdr3 cDNA encoding a human P-glycoprotein
  JOURNAL   Gene 71 (2), 401-411 (1988)
  MEDLINE   89138016
   PUBMED   2906314
COMMENT     Draft entry and computer-readable sequence for [1] kindly provided
            by P. Borst, 21-MAR-1989.
FEATURES             Location/Qualifiers
     source          1..3924
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     mRNA            <1..3924
                     /product="mdr3 mRNA"
     CDS             33..3872
                     /note="P-glycoprotein"
                     /codon_start=1
                     /protein_id="AAA36207.1"
                     /db_xref="GI:307181"
                     /translation="MDLEAAKNGTAWRPTSAEGDFELGISSKQKRKKTKTVKMIGVLT
                     LFRYSDWQDKLFMSLGTIMAIAHGSGLPLMMIVFGEMTDKFVDTAGNFSFPVNFSLSL
                     LNPGKILEEEMTRYAYYYSGLGAGVLVAAYIQVSFWTLAAGRQIRKIRQKFFHAILRQ
                     EIGWFDINDTTELNTRLTDDISKISEGIGDKVGMFFQAVATFFAGFIVGFIRGWKLTL
                     VIMAISPILGLSAAVWAKILSAFSDKELAAYAKAGAVAEEALGAIRTVIAFGGQNKEL
                     ERYQKHLENAKEIGIKKAISANISMGIAFLLIYASYALAFWYGSTLVISKEYTIGNAM
                     TVFFSILIGAFSVGQAAPCIDAFANARGAAYVIFDIIDNNPKIDSFSERGHKPDSIKG
                     NLEFNDVHFSYPSRANVKILKGLNLKVQSGQTVALVGSSGCGKSTTVQLIQRLYDPDE
                     GTINIDGQDIRNFNVNYLREIIGVVSQEPVLFSTTIAENICYGRGNVTMDEIKKAVKE
                     ANAYEFIMKLPQKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTE
                     SEAEVQAALDKAREGRTTIVIAHRLSTVRNADVIAGFEDGVIVEQGSHSELMKKEGVY
                     FKLVNMQTSGSQIQSEEFELNDEKAATRMAPNGWKSRLFRHSTQKNLKNSQMCQKSLD
                     VETDGLEANVPPVSFLKVLKLNKTEWPYFVVGTVCAIANGGLQPAFSVIFSEIIAIFG
                     PGDDAVKQQKCNIFSLIFLFLGIISFFTFFLQGFTFGKAGEILTRRLRSMAFKAMLRQ
                     DMSWFDDHKNSTGALSTRLATDAAQVQGATGTRLALIAQNIANLGTGIIISFIYGWQL
                     TLLLLAVVPIIAVSGIVEMKLLAGNAKRDKKELEAAGKIATEAIENIRTVVSLTQERK
                     FESMYVEKLYGPYRNSVQKAHIYGITFSISQAFMYFSYAGCFRFGAYLIVNGHMRFRD
                     VILVFSAIVFGAVALGHASSFAPDYAKAKLSAAHLFMLFERQPLIDSYSEEGLKPDKF
                     EGNITFNEVVFNYPTRANVPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERFYDP
                     LAGTVLLDGQEAKKLNVQWLRAQLGIVSQEPILFDCSIAENIAYGDNSRVVSQDEIVS
                     AAKAANIHPFIETLPHKYETRVGDKGTQLSGGQKQRIAIARALIRQPQILLLDEATSA
                     LDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGRVKEHGTHQQLLAQ
                     KGIYFSMVSVQAGTQNL"
BASE COUNT     1145 a    790 c    977 g   1012 t
ORIGIN

Alignment Scores:
Pred. No.:              5.8e-146      Length:         3924
Score:                  1924.50       Matches:        372
Percent Similarity:     74.10%        Conservative:   123
Best Local Similarity:  55.69%        Mismatches:     162
Query Match:            57.91%        Indels:         11
DB:                     9             Gaps:           3

US-09-873-409-1 (1-659) x HUMMDR3 (1-3924)
```

```
Qy       1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
           ::::::   |||:::|||:::|||:::|||||||||   |||   |||:::|||:::
Db    1851 GTAATTGTGGAGCAAGGAAGCCACAGCGAACTGATGAAGAAGGAAGGGGTGTACTTCAAA 1910

Qy      21 LeuVal--------------------------MetSerGlnAspIleLysLysAlaAsp 31
           ||||||                          |||::::::   :::      |||
Db    1911 CTTGTCAACATGCAGACATCAGGAAGCCAGATCCAGTCAGAAGAATTTGAACTAAATGAT 1970

Qy      32 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 51
           |||:::         :::         |||:::
Db    1971 GAAAAGGCTGCCACTAGAATGGCCCCAAATGGCTGGAAATCTCGCCTATTTAGGCATTCT 2030

Qy      52 SerValLysSerIleLys---SerAspPheIleAspLysAla---GluGluSerThrGln 69
           :::   |||::::::|||   |||         |||:::   :::         |||
Db    2031 ACTCAGAAAAACCTTAAAAATTCACAAATGTGTCAGAAGAGCCTTGATGTGGAAACCGAT 2090

Qy      70 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 89
                 |||   ::::::|||   ||||||   ||||||:::|||||||||||||||
Db    2091 GGACTTGAAGCAAATGTGCCACCAGTGTCCTTTCTGAAGGTCCTGAAACTGAATAAAACA 2150

Qy      90 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 109
           |||||||||:::   |||:::||||||:::   ::::::   ||||||   :::   |||
Db    2151 GAATGGCCCTACTTTGTCGTGGGAACAGTATGTGCCATTGCCAATGGGGGGCTTCAGCCG 2210

Qy     110 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 129
           ||||||:::||||||::::::||||||   :::||||||         |||
Db    2211 GCATTTTCAGTCATATTCTCAGAGATCATAGCGATTTTTGGACCAGGCGATGATGCAGTG 2270

Qy     130 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 149
              :::                |||:::|||:::||||||:::   ||||||:::|||   |||
Db    2271 AAGCAGCAGAAGTGCAACATATTCTCTTTGATTTTCTTATTTCTGGGAATTATTTCTTTT 2330

Qy     150 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 169
              ::::::|||:::||||||      :::|||:::||||||||||||||||||   |||
Db    2331 TTTACTTTCTTCCTTCAGGGTTTCACGTTTGGGAAAGCTGGCGAGATCCTCACCAGAAGA 2390

Qy     170 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLys 189
           ||||||   :::||||||||||||||||||   ||||||::::::||||||||:::
Db    2391 CTGCGGTCAATGGCTTTTAAAGCAATGCTAAGACAGGACATGAGCTGGTTTGATGACCAT 2450

Qy     190 GluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 209
           :::||||||||||||   |||:::|||   ||||||   |||   ||||||:::||||||
Db    2451 AAAAACAGTACTGGTGCACTTTCTACAAGACTTGCCACAGATGCTGCCCAAGTCCAAGGA 2510

Qy     210 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 229
           ||||||||||:::|||:::   ::::::   ||||||      |||:::|||      :::
Db    2511 GCCACAGGAACCAGGTTGGCTTTAATTGCACAGAATATAGCTAACCTTGGAACTGGTATT 2570

Qy     230 IleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVal 249
           ||||||||||||||||||||||||::::::|||   |||:::|||::::::   |||:::
Db    2571 ATCATATCATTTATCTACGGTTGGCAGTTAACCCTATTGCTATTAGCAGTTGTTCCAATT 2630

Qy     250 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 269
           :::||||||:::|||::::::|||         :::   |||   |||   :::||||||
Db    2631 ATTGCTGTGTCAGGAATTGTTGAAATGAAATTGTTGGCTGGAAATGCCAAAAGAGATAAA 2690
```

```
Qy     270 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIle 289
           :::||||||:::   ||||||||||||||||||||||:::|||||||||||||||:::
Db    2691 AAAGAACTGGAAGCTGCTGGAAAGATTGCAACAGAGGCAATAGAAAATATTAGGACAGTT 2750

Qy     290 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGln 309
           |||||||||||:::||||:::   ||||||   ||||||   |||   |||
Db    2751 GTGTCTTTGACCCAGGAAAGAAAATTTGAATCAATGTATGTTGAAAAATTGTATGGACCT 2810

Qy     310 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla 329
           :::||||||:::   :::||||||   |||   |||      ::::::   |||   |||
Db    2811 TACAGGAATTCTGTGCAGAAGGCACACATCTATGGAATTACTTTAGTATCTCACAAGCA 2870

Qy     330 PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGly 349
           |||:::||||||:::||||||      |||||||||||||||||||||||||   |||
Db    2871 TTTATGTATTTTTCCTATGCCGGTTGTTTTCGATTTGGTGCATATCTCATTGTGAATGGA 2930

Qy     350 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAla 369
              |||            :::   :::||||||:::||||||   :::||||||:::|||
Db    2931 CATATGCGCTTCAGAGATGTTATTCTGGTGTTTTCTGCAATTGTATTTGGTGCAGTGGCT 2990

Qy     370 IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHis 389
           :::|||               ||||||:::|||:::|||||||||   |||||||||
Db    2991 CTAGGACATGCCAGTTCATTTGCTCCAGACTATGCTAAAGCTAAGCTGTCTGCAGCCCAC 3050

Qy     390 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLys 409
           ||||||   |||   |||::::::|||   |||||||||   |||:::||||||   |||
Db    3051 TTATTCATGCTGTTTGAAAGACAACCTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAG 3110

Qy     410 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 429
           ||||||      ||||||||||:::   |||   ||||||   |||   ||||||   |||
Db    3111 CCTGATAAATTTGAAGGAAATATAACATTTAATGAAGTCGTGTTCAACTATCCCACCCGA 3170

Qy     430 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 449
              :::|||   :::|||:::||||||||||||   :::::::::|||:::|||:::|||
Db    3171 GCAAACGTGCCAGTGCTTCAGGGGCTGAGCCTGGAGGTGAAGAAAGGCCAGACACTAGCC 3230

Qy     450 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr 469
              ||||||||||||||||||||||||||||||   ||||||||||||:::|||   |||
Db    3231 CTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACGGTGGTCCAGCTCCTGGAGCGGTTCTAC 3290

Qy     470 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 489
           ||||||:::   |||   ||||||   ||||||   :::||||||:::||||||||||||
Db    3291 GACCCCTTGGCGGGGACAGTGCTTCTCGATGGTCAAGAAGCAAAGAAACTCAATGTCCAG 3350

Qy     490 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 509
           |||||||||:::|||:::   |||   ||||||||||:::||||||:::|||||||||
Db    3351 TGGCTCAGAGCTCAACTCGGAATCGTGTCTCAGGAGCCTATCCTATTTGACTGCAGCATT 3410

Qy     510 AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu 529
           ||||||||||||||||||||||||||||||||||||            |||||||||
Db    3411 GCCGAGAATATTGCCTATGGAGACAACAGCCGGGTTGTATCACAGGATGAAATTGTGAGT 3470

Qy     530 AlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThr 549
           ||||||   |||||||||||||||   |||||||||   ||||||   ||||||   |||
Db    3471 GCAGCCAAAGCTGCCAACATACATCCTTTCATCGAGACGTTACCCCACAAATATGAAACA 3530

Qy     550 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAla 569
           :::||||||   ||||||   |||||||||||||||||||||||||:::|||||||||
Db    3531 AGAGTGGGAGATAAGGGGACTCAGCTCTCAGGAGGTCAAAAACAGAGGATTGCTATTGCC 3590

Qy     570 ArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsp 589
           |||||||||:::::::::|||:::||||||||||||||||||||||||||||||||
Db    3591 CGAGCCCTCATCAGACAACCTCAAATCCTCCTGTTGGATGAAGCTACATCAGCTCTGGAT 3650

Qy     590 AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys 609
           :::||||||||||||||||||         ||||||||||||||||   ||||||||||
Db    3651 ACTGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGC 3710

Qy     610 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 629
           :::|||:::   ||||||||||||   ||||||||||||||||||||||||
Db    3711 ATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAG 3770

Qy     630 AsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyr 649
           ||||||::::::||||||   ||||||||||||:::||||||      :::   ||||||
Db    3771 AATGGGAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTAT 3830

Qy     650 PheLysLeuValAsnAlaGlnSer 657
           |||   :::|||:::   |||:::
Db    3831 TTTTCAATGGTCAGTGTCCAGGCT 3854
```

```
RESULT 10
CRUPGP1185
LOCUS       CRUPGP1185            3987 bp    mRNA    linear   ROD 27-APR-1993
DEFINITION  Chinese hamster p-glycoprotein mRNA (clone ADX185), complete cds.
ACCESSION   M59254
VERSION     M59254.1  GI:191156
KEYWORDS    multidrug resistance glycoprotein; p-glycoprotein.
SOURCE      Chinese hamster, cDNA to mRNA, clone ADX185.
  ORGANISM  Cricetulus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
            Cricetulus.
REFERENCE   1  (bases 1 to 3987)
  AUTHORS   Devine,S.E., Hussain,A., Davide,J.P. and Melera,P.W.
  TITLE     Full length and alternatively spliced pgp1 transcripts in
            multidrug-resistant Chinese hamster lung cells
  JOURNAL   J. Biol. Chem. 266 (7), 4545-4555 (1991)
  MEDLINE   91154265
   PUBMED   1671863
FEATURES             Location/Qualifiers
     source          1..3987
                     /organism="Cricetulus sp."
                     /db_xref="taxon:10031"
                     /cell_type="fibroblast"
                     /tissue_type="lung"
     gene            1..3987
                     /gene="pgp-1"
     CDS             115..3624
                     /gene="pgp-1"
                     /codon_start=1
                     /product="p-glycoprotein"
                     /protein_id="AAA37005.1"
                     /db_xref="GI:191157"
                     /translation="MTTYAYYYTGIGAGVLIVAYIQVSFWCLAAGRQIHKIRQKFFHA
                     IMNQEIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQAMATFFGGFIIGFTRGW
                     KLTLVILAISPVLGLSAGIWAKILSSFTDKELQAYAKAGAVAEEVLAAIRTVIAFGGQ
                     KKELERYNNNLEEAKRLGIKKAITANISMGAAFLLIYASYALAFWYGTSLVISKEYSI
                     GQVLTVFFAVLIAPFSIGQASPNIEAFANARGAAYEIFNIIDNKPSIDSFSKNGYKPD
                     NIKGNLEFKNIHFSYPSRKDVQILKGLNLKVQSGQTVALVGNSGCGKSTTVQLLQRLY
                     DPTEGVVSIDGQDIRTINVRYLREIIGVVSQEPVLFATTIAENIRYGRENVTMDEIEK
                     AVKEANAYDFIMKLPHKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSA
                     LDTESEAVVQAALDKAREGRTTIVIAHRLSTVRNADIIAGFDGGVIVEQGNHEELMRE
                     KGIYFKLVMTQTAGNEIELGNEVGESKNEIDNLDMSSKDSASSLIRRRSTRRSIRGPH
                     DQDRKLSTKEALDEDVPPISFWRILKLNSSEWPYFVVGIFCAIVNGALQPAFSIIFSK
                     VVGVFTRNTDDETKRHDSNLFSLLFLILGVISFITFFLQGFTFGKAGEILTKRLRYMV
                     FKSMLRQDVSWFDNPKNTTGALTTRLANDAGQVKGATGARLAVITQNIANLGTGIIIS
                     LIYGWQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGSGKIATEAIENFRTVV
                     SLTREQKFENMYAQSLQIPYRNALKKAHVFGITFSFTQAMMYFSYAACFRFGAYLVAR
                     ELMTFENVLLVFSAIVFGAMAVGQVSSFAPDYAKAKVSASHIIMIIEKVPSIDSYSTG
                     GLKPNTLEGNVKFNEVVFNYPTRPDIPVLQGLNLEVKKGQTLALVGSSGCGKSTVVQL
                     LERFYDPMAGTVFLDGKEVNQLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVV
                     SQDEIERAAKEANIHQFIESLPDKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILL
                     LDEATSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVIQNGKVKEHGT
                     HQQLLAQKGIYFSMVSVQAGAKR"
BASE COUNT     1194 a    822 c    944 g   1027 t
ORIGIN      Map position 1q26.

Alignment Scores:
Pred. No.:              2.88e-145     Length:         3987
Score:                  1916.00       Matches:        360
Percent Similarity:     74.96%        Conservative:   143
Best Local Similarity:  53.65%        Mismatches:     154
Query Match:            57.66%        Indels:         14
DB:                     10            Gaps:           4
```

```
US-09-873-409-1 (1-659) x CRUPGP1185 (1-3987)
Qy       1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
             ::::::    |||:::|||    |||   ||||||||   ::::::|||:::|||:::
Db    1597 GTCATTGTGGAGCAAGGAAATCATGAAGAGCTCATGAGAGAGAAGGGCATTTACTTCAAA 1656

Qy      21 LeuValMetSerGln-----------------------------AspIleLysLysAla 30
             |||||||||:::|||                             ::::::   ::::::
Db    1657 CTTGTCATGACACAGACAGCAGGAAATGAAATTGAATTAGGAAATGAAGTTGGTGAGTCT 1716

Qy      31 AspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu------ 48
             :::::::::::::::           |||:::::::         :::||||||
Db    1717 AAAAATGAAATTGATAATTTAGACATGTCTTCAAAAGATTCAGCATCCAGTCTAATCAGA 1776

Qy      49 ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr 68
                     :::   :::||||||:::                    :::   ||||||
Db    1777 AGGAGATCAACTCGCAGAAGTATCCGTGGACCACATGACCAAGACAGGAAGCTTAGTACC 1836

Qy      69 ---GlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
              ::::::   :::      :::|||   :::|||      :::||||||||||||||||
Db    1837 AAAGAGGCCTTGGATGAAGATGTACCTCCAATTTCCTTTTGGCGTATCCTAAAGTTGAAT 1896

Qy      88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
                   ||||||||||:::   |||:::|||         :::::::::||||||   :::
Db    1897 TCATCTGAATGGCCATATTTGTGGTTGGTATATTCTGTGCCATAGTAAATGGAGCCTTG 1956

Qy     108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAsp 126
             |||   ||||||||||||||||:::|||::::::                |||   |||
Db    1957 CAACCAGCATTCTCAATAATATTTTCTAAGGTTGTAGGGGTTTTCACAAGAAATACTGAT 2016

Qy     127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
                   |||   :::||||||:::   ::::::|||::::::|||:::||||||||||||
Db    2017 GATGAAACCAAACGACATGATAGCAACTTGTTTTCCTTATTGTTTCTGATCCTTGGGGTC 2076

Qy     147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
             |||   |||:::::::::|||:::||||||      :::|||:::||||||||||||||||
Db    2077 ATTTCTTTTATTACATTTTTTCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTC 2136

Qy     167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
             |||   |||||||||::::::   ||||||:::||||||   ||||||::::::||||||
Db    2137 ACCAAGCGACTCCGATACATGGTTTTCAAATCCATGCTGAGACAGGATGTGAGCTGGTTT 2196

Qy     187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
             |||      :::|||:::|||||||||   |||||||||   ||||||   |||   |||
Db    2197 GATAACCCTAAAAACACCACTGGAGCATTGACCACCAGGCTCGCCAACGATGCTGGTCAA 2256

Qy     207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
             ::::::|||||||||||||:::|||:::   |||:::||||||||      |||:::|||
Db    2257 GTTAAAGGGGCTACAGGAGCCAGACTTGCTGTCATTACCCAGAATATAGCAAATCTTGGG 2316

Qy     227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
                   :::|||||||||   ||||||||||||::::::|||   |||:::|||:::|||
Db    2317 ACAGGAATCATCATATCCCTAATCTATGGCTGGCAGTTGACACTTCTACTCTTAGCAATT 2376

Qy     247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
                |||:::::|||:::   |||::::::|||         ::::::|||   |||
Db    2377 GTCCCCATCATTGCAATAGCAGGAGTGGTTGAGATGAAAATGTTGTCTGGACAAGCACTA 2436

Qy     267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
             |||||||||:::||||||:::   :::||||||||||||||||||||:::||||||
Db    2437 AAAGATAAGAAGGAGCTAGAAGGTTCTGGGAAGATTGCTACTGAAGCAATAGAGAACTTC 2496

Qy     287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
             ||||||:::|||||||||||||||:::   ||||||   ||||||   :::   |||
Db    2497 CGCACTGTCGTCTCTTTGACTCGGGAGCAGAAGTTTGAAAATATGTATGCCCAGAGCTTG 2556

Qy     307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
             |||      :::||||||      |||||||||   :::   |||      ::::::|||
Db    2557 CAGATACCATACAGAAATGCTCTGAAGAAAGCACACGTCTTTGGGATCACATTCTCCTTC 2616

Qy     327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
             :::   |||   :::||||||:::|||||||||   ||||||||||||||||||||:::
Db    2617 ACCCAGGCCATGATGTATTTTTCGTATGCTGCTTGTTTCCGGTTTGGTGCTTACTTGGTG 2676

Qy     347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
                         ||||||   |||   :::   :::||||||:::||||||   :::|||
Db    2677 GCACGGGAATTAATGACATTTGAAAATGTTCTATTAGTATTCTCAGCTATTGTCTTTGGT 2736

Qy     367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
             ||||||||:::|||:::                  ||||||:::|||:::||||||||
Db    2737 GCCATGGCAGTTGGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTGTCG 2796

Qy     387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
             |||:::|||:::         ::::::||||||   |||:::|||||||||   |||
Db    2797 GCATCGCATATCATCATGATCATTGAAAAAGTCCCTTCCATTGACAGCTACAGCACCGGA 2856

Qy     407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
             |||   ||||||:::|||   |||||||||::::::   ||||||   |||   |||
Db    2857 GGCCTGAAGCCTAATACATTGGAAGGAAATGTGAAATTTAATGAAGTCGTGTTCAACTAC 2916

Qy     427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
             |||   ||||||||:::   :::|||:::||||||:::|||      :::::::::|||:::
Db    2917 CCCACCCGGCCTGACATCCCAGTGCTTCAGGGGCTGAACCTGGAGGTGAAGAAGGGCCAG 2976

Qy     447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
             |||:::|||   |||||||||||||||||||||||||||   |||||||||||:::
Db    2977 ACGCTGGCCCTGGTGGGCAGCAGTGGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAG 3036

Qy     467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
             |||   ||||||||:::   |||   |||      ||||||   :::      :::|||
Db    3037 CGGTTCTACGACCCCATGGCCGGAACAGTGTTTCTAGATGGCAAAGAAGTAAACCAACTA 3096

Qy     487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
             ||||||||||||||||||:::   :::      ||||||||||:::||||||:::
Db    3097 AATGTCCAGTGGCTCCGAGCACACCTAGGCATTGTGTCCCAAGAGCCCATCCTGTTTGAC 3156

Qy     507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
             |||||||||||||||||||||||||||||||||||||||||||||         ||||||
Db    3157 TGCAGCATCGCTGAGAACATCGCTTATGGAGACAACAGCCGGGTCGTGTCCCAGGATGAG 3216

Qy     527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
             |||:::   ||||||   ||||||||||||   |||||||||   ||||||:::|||
Db    3217 ATCGAGAGGGCGGCGAAGGAAGCCAACATCCACCAGTTCATTGAGTCACTGCCTGACAAA 3276

Qy     547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
             ||||||||:::||||||   ||||||   ||||||||||||||||||||||||:::
Db    3277 TACAACACCAGAGTAGGAGACAAAGGGACTCAGCTGTCGGGTGGGCAGAAACAACGCATC 3336

Qy     567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
             ||||||||||||||||||:::::::::||||||||||||||||||||||||
Db    3337 GCCATAGCACGCGCCCTCGTCAGACAGCCTCACATTTTGCTCCTGGATGAAGCAACATCA 3396

Qy     587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
             |||||||||      :::|||||||||||||||   ||||||||||||||||||   |||
Db    3397 GCCCTGGATACAGAGAGTGAAAAGGTTGTCCAGGAAGCTCTGGACAAAGCCAGAGAAGGC 3456

Qy     607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
             |||||||||:::|||:::   ||||||||||||   ||||||||||||||||||||||||
Db    3457 CGCACCTGCATTGTGATCGCTCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTG 3516

Qy     627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
             |||:::   ||||||||||:::||||||   ||||||||||||:::||||||      :::
Db    3517 GTGATTCAGAATGGCAAGGTCAAGGAGCATGGCACCCACCAGCAGCTGCTGGCACAGAAA 3576

Qy     647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
                |||||||||   :::|||:::   |||:::
Db    3577 GGCATCTATTTCTCCATGGTCAGTGTGCAGGCT 3609
```

RESULT 11
CRUPGPI
LOCUS CRUPGPI 4296 bp mRNA linear ROD 22-JUN-1995
DEFINITION C.griseus P-glycoprotein (isoform class I) mRNA, complete cds.
ACCESSION M60040
VERSION M60040.1 GI:191164
KEYWORDS multidrug resistance; p-glycoprotein; transmembrane protein.
SOURCE C.griseus adult liver and ovary, cDNA to mRNA.

```
  ORGANISM  Cricetulus griseus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
            Cricetulus.
REFERENCE   1  (bases 1 to 4296)
  AUTHORS   Endicott,J.A., Sarangi,F. and Ling,V.
  TITLE     Complete cDNA sequences encoding the Chinese hamster P-glycoprotein
            gene family
  JOURNAL   DNA Seq. 2 (2), 89-101 (1991)
  MEDLINE   92135896
   PUBMED   1685679
FEATURES             Location/Qualifiers
     source          1..4296
                     /organism="Cricetulus griseus"
                     /db_xref="taxon:10029"
                     /tissue_type="liver and ovary"
                     /dev_stage="adult"
     CDS             103..3933
                     /note="putative"
                     /codon_start=1
                     /product="p-glycoprotein isoform I"
                     /protein_id="AAA68883.1"
                     /db_xref="GI:191165"
                     /translation="MEFEEDFSGRKDKNFLKMGRKSKKEKKEKKPVVSVFTMFRYAGW
                     LDRLYMLVGTLAAIIHGVALPLMMLVFGDMTDSFASVGNIPTNATNNATQVNASDIFG
                     KLEEEMTTYAYYYTGIGAGVLIVAYIQVSFWCLAAGRQIHKIRQKFFHAIMNQEIGWF
                     DVHDVGELNTRLTDDVSKINEGIGDKIGMFFQAMATFFGGFIIGFTRGWKLTLVILAI
                     SPVLGLSAGIWAKILSSFTDKELQAYAKAGAVAEEVLAAIRTVIAFGGQKKELERYNN
                     NLEEAKRLGIKKAITANISMGAAFLLIYASYALAFWYGTSLVISKEYSIGQVLTVFFA
                     VLIGAFSIGQASPNIEAFANARGAAYEIFNIIDNKPSIDSFSKNGYKPDNIKGNLEFK
                     NIHFSYPSRKDVQILKGLNLKVQSGQTVALVGNSGCGKSTTVQLLQRLYDPTEGVVSI
                     DGQDIRTINVRYLREIIGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKEANAYD
                     FIMKLPHKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTESEAVV
                     QAALDKAREGRTTIVIAHRLSTVRNADIIAGFDGGVIVEQGNHEELMREKGIYFKLVM
                     TQTAGNEIELGNEVGESKNEIDNLDMSSKDSASSLIRRRSTRRSIRGPHDQDRKLSTK
                     EALDEDVPPISFWRILKLNSSEWPYFVVGIFCAIVNGALQPAFSIIFSKVVGVFTRNT
                     DDETKRHDSNLFSLLFLILGVISFITFFLQGFTFGKAGEILTKRLRYMVFKSMLRQDV
                     SWFDNPKNTTGALTTRLANDAGQVKGATGARLAVITQNIANLGTGIIISLIYGWQLTL
                     LLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGSGKIATEAIENFRTVVSLTREQKFE
                     NMYAQSLQIPYRNALKKAHVFGITFSFTQAMMYFSYAACFRFGAYLVARELMTFENVL
                     LVFSAIVFGAMAVGQVSSFAPDYAKAKVSASHIIMIIEKVPSIDSYSTGGLKPNTLEG
                     NVKFNEVVFNYPTRPDIPVLQGLNLEVKKGQTLALVGSSGCGKSTVVQLLERFYDPMA
                     GTVFLDGKEVNQLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSQDEIERAA
                     KEANIHQFIESLPDKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEATSALD
                     TESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVIQNGKVKEHGTHQQLLAQKG
                     IYFSMVSVQAGAKR"
BASE COUNT     1271 a    892 c   1029 g   1104 t
ORIGIN

Alignment Scores:
Pred. No.:              3.19e-145      Length:         4296
Score:                  1916.00        Matches:        360
Percent Similarity:     74.96%         Conservative:   143
Best Local Similarity:  53.65%         Mismatches:     154
Query Match:            57.66%         Indels:         14
DB:                     10             Gaps:           4

US-09-873-409-1 (1-659) x CRUPGPI (1-4296)

Qy        1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
              ::::::   |||:::|||   |||   ||||||||    ::::::|||:::|||:::
Db     1906 GTCATTGTGGAGCAAGGAAATCATGAAGAGCTCATGAGAGAGAAGGGCATTTACTTCAAA 1965

Qy       21 LeuValMetSerGln------------------------------AspIleLysLysAla 30
            |||||||||:::|||                              ::::::   ::::::
Db     1966 CTTGTCATGACACAGACAGCAGGAAATGAAATTGAATTAGGAAATGAAGTTGGTGAGTCT 2025

Qy       31 AspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu------ 48
                  ::::::::::::::         |||::::::        :::||||||
Db     2026 AAAAATGAAATTGATAATTTAGACATGTCTTCAAAAGATTCAGCATCCAGTCTAATCAGA 2085

Qy       49 ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr 68
                     :::   :::||||||:::                    :::   ||||||
Db     2086 AGGAGATCAACTCGCAGAAGTATCCGTGGACCACATGACCAAGACAGGAAGCTTAGTACC 2145

Qy       69 ---GlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
               ::::::   :::      :::|||   :::|||      :::||||||||||||||
Db     2146 AAAGAGGCCTTGGATGAAGATGTACCTCCAATTTCCTTTTGGCGTATCCTAAAGTTGAAT 2205

Qy       88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
                  |||||||||:::   |||:::|||         :::::::::||||||   :::
Db     2206 TCATCTGAATGGCCATATTTTGTGGTTGGTATATTCTGTGCCATAGTAAATGGAGCCTTG 2265

Qy      108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAsp 126
              |||   ||||||||||||||:::|||::::::   :::|||          |||   |||
Db     2266 CAACCAGCATTCTCAATAATATTTTCTAAGGTTGTAGGGGTTTTCACAAGAAATACTGAT 2325

Qy      127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
                  |||   :::||||||:::   ::::::|||::::::|||:::||||||||||||
Db     2326 GATGAAACCAAACGACATGATAGCAACTTGTTTTCCTTATTGTTTCTGATCCTTGGGGTC 2385

Qy      147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
            |||   |||:::::::::|||:::||||||         :::|||:::|||||||||||||
Db     2386 ATTTCTTTTATTACATTTTTTCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTC 2445

Qy      167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
            |||   ||||||||||::::::   ||||||:::||||||   ||||||::::::||||||
Db     2446 ACCAAGCGACTCCGATACATGGTTTTCAAATCCATGCTGAGACAGGATGTGAGCTGGTTT 2505

Qy      187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
            |||      :::|||:::|||||||||   ||||||||||||   |||      |||
Db     2506 GATAACCCTAAAAACACCACTGGAGCATTGACCACCAGGCTCGCCAACGATGCTGGTCAA 2565

Qy      207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
            ::::::|||||||||||||||:::|||:::      |||:::|||||||||      |||:::|||
Db     2566 GTTAAAGGGGCTACAGGAGCCAGACTTGCTGTCATTACCCAGAATATAGCAAATCTTGGG 2625

Qy      227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
               :::|||||||||   ||||||||||||:::::::::|||   |||:::|||:::|||
Db     2626 ACAGGAATCATCATATCCCTAATCTATGGCTGGCAGTTGACACTTCTACTCTTAGCAATT 2685

Qy      247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
               |||::::::|||:::   |||::::::|||         ::::::|||   |||
Db     2686 GTCCCCATCATTGCAATAGCAGGAGTGGTTGAGATGAAAATGTTGTCTGGACAAGCACTA 2745

Qy      267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
            ||||||||||||:::|||:::   :::||||||||||||||||||:::||||||
Db     2746 AAAGATAAGAAGGAGCTAGAAGGTTCTGGGAAGATTGCTACTGAAGCAATAGAGAACTTC 2805

Qy      287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
            ||||||:::||||||||||||||||||:::   ||||||   ||||||   :::   |||
Db     2806 CGCACTGTCGTCTCTTTGACTCGGGAGCAGAAGTTTGAAAATATGTATGCCCAGAGCTTG 2865

Qy      307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
            |||      :::||||||      ||||||||||   :::   |||      ::::::|||
Db     2866 CAGATACCATACAGAAATGCTCTGAAGAAAGCACACGTCTTTGGGATCACATTCTCCTTC 2925

Qy      327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
            :::   |||   :::|||||||:::|||||||||   ||||||||||||||||||||:::
Db     2926 ACCCAGGCCATGATGTATTTTTCGTATGCTGCTTGTTTCCGGTTTGGTGCTTACTTGGTG 2985

Qy      347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
                     ||||||   |||   :::   :::||||||:::||||||     :::|||
Db     2986 GCACGGGAATTAATGACATTTGAAAATGTTCTATTAGTATTCTCAGCTATTGTCTTTGGT 3045

Qy      367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
            ||||||||:::|||:::               ||||||:::|||:::|||||||||
Db     3046 GCCATGGCAGTTGGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTGTCG 3105

Qy      387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
            |||:::|||:::      ::::::||||||   |||:::|||||||||   |||
Db     3106 GCATCGCATATCATCATGATCATTGAAAAAGTCCCTTCCATTGACAGCTACAGCACCGGA 3165

Qy      407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
            |||   ||||||:::|||   ||||||||||:::::::|||   ||||||   |||   |||
Db     3166 GGCCTGAAGCCTAATACATTGGAAGGAAATGTGAAATTTAATGAAGTCGTGTTCAACTAC 3225
```

```
Qy     427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
           |||   ||||||||:::   :::|||:::||||||:::|||   :::::::::|||:::
Db    3226 CCCACCCGGCCTGACATCCCAGTGCTTCAGGGGCTGAACCTGGAGGTGAAGAAGGGCCAG 3285

Qy     447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
           |||:::|||   ||||||||||||||||||||||||||||||   ||||||||||||:::
Db    3286 ACGCTGGCCCTGGTGGGCAGCAGTGGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAG 3345

Qy     467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
           |||   |||||||||:::   |||   |||      ||||||   :::      :::|||
Db    3346 CGGTTCTACGACCCCATGGCCGGAACAGTGTTTCTAGATGGCAAAGAAGTAAACCAACTA 3405

Qy     487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
           ||||||||||||||||||:::   :::   ||||||   |||||||||:::||||||:::
Db    3406 AATGTCCAGTGGCTCCGAGCACACCTAGGCATTGTGTCCCAAGAGCCCATCCTGTTTGAC 3465

Qy     507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
           ||||||||||||||||||||||||||||||||||||||||||||||||      ||||||
Db    3466 TGCAGCATCGCTGAGAACATCGCTTATGGAGACAACAGCCGGGTCGTGTCCCAGGATGAG 3525

Qy     527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
           |||:::   ||||||      ||||||||||||   |||||||||   ||||||:::|||
Db    3526 ATCGAGAGGGCGGCGAAGGAAGCCAACATCCACCAGTTCATTGAGTCACTGCCTGACAAA 3585

Qy     547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
           |||||||||:::||||||   ||||||   |||||||||||||||||||||||||||:::
Db    3586 TACAACACCAGAGTAGGAGACAAAGGGACTCAGCTGTCGGGTGGGCAGAAACAACGCATC 3645

Qy     567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
           ||||||||||||||||||:::::::::|||   ||||||||||||||||||||||||||||
Db    3646 GCCATAGCACGCGCCCTCGTCAGACAGCCTCACATTTTGCTCCTGGATGAAGCAACATCA 3705

Qy     587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
           |||||||||   :::||||||||||||||||||||||||||||||||||||   |||
Db    3706 GCCCTGGATACAGAGAGTGAAAAGGTTGTCCAGGAAGCTCTGGACAAAGCCAGAGAAGGC 3765

Qy     607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
           |||||||||:::|||:::   ||||||||||||   ||||||||||||||||||||||||
Db    3766 CGCACCTGCATTGTGATCGCTCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTG 3825

Qy     627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
           |||:::   |||||||||:::||||||   ||||||||||||:::||||||      :::
Db    3826 GTGATTCAGAATGGCAAGGTCAAGGAGCATGGCACCCACCAGCAGCTGCTGGCACAGAAA 3885

Qy     647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
              |||||||||   :::|||:::   |||:::
Db    3886 GGCATCTATTTCTCCATGGTCAGTGTGCAGGCT 3918
```

```
RESULT 12
CRUPGP1165
LOCUS       CRUPGP1165            4304 bp    mRNA    linear   ROD 27-APR-1993
DEFINITION  Chinese hamster p-glycoprotein mRNA (clone ADX165), complete cds.
ACCESSION   M59253
VERSION     M59253.1  GI:191154
KEYWORDS    multidrug resistance glycoprotein; p-glycoprotein.
SOURCE      Chinese hamster, cDNA to mRNA, clone ADX165.
  ORGANISM  Cricetulus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
            Cricetulus.
REFERENCE   1  (bases 1 to 4304)
  AUTHORS   Devine,S.E., Hussain,A., Davide,J.P. and Melera,P.W.
  TITLE     Full length and alternatively spliced pgp1 transcripts in
            multidrug-resistant Chinese hamster lung cells
  JOURNAL   J. Biol. Chem. 266 (7), 4545-4555 (1991)
  MEDLINE   91154265
   PUBMED   1671863
FEATURES             Location/Qualifiers
     source          1. .4304
                     /organism="Cricetulus sp."
                     /db_xref="taxon:10031"
                     /cell_type="fibroblast"
                     /tissue_type="lung"
     gene            1. .4304
                     /gene="pgp-1"
     CDS             111. .3941
                     /gene="pgp-1"
                     /codon_start=1
                     /product="p-glycoprotein"
                     /protein_id="AAA37004.1"
                     /db_xref="GI:191155"
                     /translation="MEFEEDFSGRKDKNFLKMGRKSKKEKKEKKPVVSVFTMFRYAGW
                     LDRLYMLVGTLAAIIHGVALPLMMLVFGDMTDSFASVGNIPTNATNNATQVNASDIFG
                     KLEEEMTTYAYYYTGIGAGVLIVAYIQVSFWCLAAGRQIHKIRQKFFHAIMNQEIGWF
                     DVHDVGELNTRLTDDVSKINEGIGDKIGMFFQAMATFFGGFIIGFTRGWKLTLVILAI
                     SPVLGLSAGIWAKILSSFTDKELQAYAKAGAVAEEVLAAIRTVIAFGGQKKELERYNN
                     NLEEAKRLGIKKAITANISMGAAFLLIYASYALAFWYGTSLVISKEYSIGQVLTVFFA
                     VLIAPFSIGQASPNIEAFANARGAAYEIFNIIDNKPSIDSFSKNGYKPDNIKGNLEFK
                     NIHFSYPSRKDVQILKGLNLKVQSGQTVALVGNSGCGKSTTVQLLQRLYDPTEGVVSI
                     DGQDIRTINVRYLREIIGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKEANAYD
                     FIMKLPHKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTESEAVV
                     QAALDKAREGRTTIVIAHRLSTVRNADIIAGFDGGVIVEQGNHEELMREKGIYFKLVM
                     TQTAGNEIELGNEVGESKNEIDNLDMSSKDSASSLIRRRSTRRSIRGPHDQDRKLSTK
                     EALDEDVPPISFWRILKLNSSEWPYFVVGIFCAIVNGALQPAFSIIFSKVVGVFTRNT
                     DDETKRHDSNLFSLLFLILGVISFITFFLQGFTFGKAGEILTKRLRYMVFKSMLRQDV
                     SWFDNPKNTTGALTTRLANDAGQVKGATGARLAVITQNIANLGTGIIISLIYGWQLTL
                     LLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGSGKIATEAIENFRTVVSLTREQKFE
                     NMYAQSLQIPYRNALKKAHVFGITFSFTQAMMYFSYAACFRFGAYLVARELMTFENVL
                     LVFSAIVFGAMAVGQVSSFAPDYAKAKVSASHIIMIIEKVPSIDSYSTGGLKPNTLEG
                     NVKFNEVVFNYPTRPDIPVLQGLNLEVKKGQTLALVGSSGCGKSTVVQLLERFYDPMA
                     GTVFLDGKEVNQLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSQDEIERAA
                     KEANIHQFIESLPDKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEATSALD
                     TESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVIQNGKVKEHGTHQQLLAQKG
                     IYFSMVSVQAGAKR"
BASE COUNT     1275 a     895 c    1031 g    1103 t
ORIGIN      Map position 1q26.

Alignment Scores:
Pred. No.:              3.19e-145      Length:         4304
Score:                  1916.00        Matches:        360
Percent Similarity:     74.96%         Conservative:   143
Best Local Similarity:  53.65%         Mismatches:     154
Query Match:            57.66%         Indels:         14
DB:                     10             Gaps:           4

US-09-873-409-1 (1-659) x CRUPGP1165 (1-4304)
```

```
Qy       1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
             ::::::   |||:::|||   |||   |||||||||   ::::::|||:::|||:::
Db    1914 GTCATTGTGGAGCAAGGAAATCATGAAGAGCTCATGAGAGAGAAGGGCATTTACTTCAAA 1973

Qy      21 LeuValMetSerGln---------------------------AspIleLysLysAla 30
           |||||||||:::|||                            ::::::   ::::::
Db    1974 CTTGTCATGACACAGACAGCAGGAAATGAAATTGAATTAGGAAATGAAGTTGGTGAGTCT 2033

Qy      31 AspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu------ 48
              :::::::::::::::         |||::::::         :::||||||
Db    2034 AAAAATGAAATTGATAATTTAGACATGTCTTCAAAAGATTCAGCATCCAGTCTAATCAGA 2093

Qy      49 ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr 68
              :::   :::||||||:::                     :::   ||||||
Db    2094 AGGAGATCAACTCGCAGAAGTATCCGTGGACCACATGACCAAGACAGGAAGCTTAGTACC 2153

Qy      69 ---GlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
              ::::::   :::      :::|||   :::|||      :::||||||||||||||
Db    2154 AAAGAGGCCTTGGATGAAGATGTACCTCCAATTTCCTTTTGGCGTATCCTAAAGTTGAAT 2213

Qy      88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
                 ||||||||||:::   |||:::|||         :::::::::||||||   :::
Db    2214 TCATCTGAATGGCCATATTTTGTGGTTGGTATATTCTGTGCCATAGTAAATGGAGCCTTG 2273

Qy     108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAsp 126
              |||   ||||||||||||||:::|||::::::      :::|||      |||   |||
Db    2274 CAACCAGCATTCTCAATAATATTTTCTAAGGTTGTAGGGGTTTTCACAAGAAATACTGAT 2333
```

```
Qy     127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
               |||    :::|||||:::    ::::::|||::::::|||:::|||||||||||||
Db    2334 GATGAAACCAAACGACATGATAGCAACTTGTTTTCCTTATTGTTTCTGATCCTTGGGGTC 2393

Qy     147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
           |||   |||:::::::::|||:::||||||       :::|||:::||||||||||||||
Db    2394 ATTTCTTTTATTACATTTTTTCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTC 2453

Qy     167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
           |||   |||||||||::::::   ||||||:::||||||   ||||||::::::||||||
Db    2454 ACCAAGCGACTCCGATACATGGTTTTCAAATCCATGCTGAGACAGGATGTGAGCTGGTTT 2513

Qy     187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
           |||         :::|||:::||||||   |||||||||   ||||||   |||      |||
Db    2514 GATAACCCTAAAAACACCACTGGAGCATTGACCACCAGGCTCGCCAACGATGCTGGTCAA 2573

Qy     207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
           ::::::||||||||||||||||:::|||:::   |||:::|||||||||      |||:::|||
Db    2574 GTTAAAGGGGCTACAGGAGCCAGACTTGCTGTCATTACCCAGAATATAGCAAATCTTGGG 2633

Qy     227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
               :::|||||||||   ||||||||||||::::::|||   |||:::|||:::|||
Db    2634 ACAGGAATCATCATATCCCTAATCTATGGCTGGCAGTTGACACTTCTACTCTTAGCAATT 2693

Qy     247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
                 |||::::::|||:::   |||::::::|||         ::::::|||   |||
Db    2694 GTCCCCATCATTGCAATAGCAGGAGTGGTTGAGATGAAAATGTTGTCTGGACAAGCACTA 2753

Qy     267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
           |||||||||:::||||||:::   :::||||||||||||||||||||||:::||||||
Db    2754 AAAGATAAGAAGGAGCTAGAAGGTTCTGGGAAGATTGCTACTGAAGCAATAGAGAACTTC 2813

Qy     287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
           ||||||:::||||||||||||||||:::   ||||||   ||||||      :::   |||
Db    2814 CGCACTGTCGTCTCTTTGACTCGGGAGCAGAAGTTTGAAAATATGTATGCCCAGAGCTTG 2873

Qy     307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
           |||      :::||||||      ||||||||||   :::   |||         ::::::|||
Db    2874 CAGATACCATACAGAAATGCTCTGAAGAAAGCACACGTCTTTGGGATCACATTCTCCTTC 2933

Qy     327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
           :::   |||   :::||||||:::||||||||   ||||||||||||||||||||||:::
Db    2934 ACCCAGGCCATGATGTATTTTTCGTATGCTGCTTGTTTCCGGTTTGGTGCTTACTTGGTG 2993

Qy     347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
                       ||||||   |||   :::   :::||||||:::||||||   :::|||
Db    2994 GCACGGGAATTAATGACATTTGAAAATGTTCTATTAGTATTCTCAGCTATTGTCTTTGGT 3053

Qy     367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
           |||||||||:::|||:::                  ||||||:::|||:::|||||||||
Db    3054 GCCATGGCAGTTGGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTGTCG 3113

Qy     387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
           |||:::|||:::         ::::::||||||   |||:::||||||||||   |||
Db    3114 GCATCGCATATCATCATGATCATTGAAAAAGTCCCTTCCATTGACAGCTACAGCACCGGA 3173

Qy     407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
           |||   ||||||:::|||   |||||||||:::::::|||   ||||||   |||   |||
Db    3174 GGCCTGAAGCCTAATACATTGGAAGGAAATGTGAAATTTAATGAAGTCGTGTTCAACTAC 3233

Qy     427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
           |||   |||||||||:::   :::|||:::||||||:::|||   :::::::::|||:::
Db    3234 CCCACCCGGCCTGACATCCCAGTGCTTCAGGGGCTGAACCTGGAGGTGAAGAAGGGCCAG 3293

Qy     447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
           |||:::|||   ||||||||||||||||||||||||||||   ||||||||||||:::
Db    3294 ACGCTGGCCCTGGTGGGCAGCAGTGGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAG 3353

Qy     467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
           |||   |||||||||:::   |||   |||      ||||||      :::      :::|||
Db    3354 CGGTTCTACGACCCCATGGCCGGAACAGTGTTTCTAGATGGCAAAGAAGTAAACCAACTA 3413

Qy     487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
           ||||||||||||||||||:::   :::   ||||||   ||||||||:::||||||:::
Db    3414 AATGTCCAGTGGCTCCGAGCACACCTAGGCATTGTGTCCCAAGAGCCCATCCTGTTTGAC 3473

Qy     507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
           |||||||||||||||||||||||||||||||||||||||||||||         ||||||
Db    3474 TGCAGCATCGCTGAGAACATCGCTTATGGAGACAACAGCCGGGTCGTGTCCCAGGATGAG 3533

Qy     527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
           |||:::   ||||||      ||||||||||||   |||||||||   ||||||:::|||
Db    3534 ATCGAGAGGGCGGCGAAGGAAGCCAACATCCACCAGTTCATTGAGTCACTGCCTGACAAA 3593

Qy     547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
           |||||||||:::||||||   ||||||   ||||||||||||||||||||||||||||:::
Db    3594 TACAACACCAGAGTAGGAGACAAAGGGACTCAGCTGTCGGGTGGGCAGAAACAACGCATC 3653

Qy     567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
           |||||||||||||||||||:::::::::|||   ||||||||||||||||||||||||
Db    3654 GCCATAGCACGCGCCCTCGTCAGACAGCCTCACATTTTGCTCCTGGATGAAGCAACATCA 3713

Qy     587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
           |||||||    :::|||||||||||||||||||   ||||||||||||||||||   |||
Db    3714 GCCCTGGATACAGAGAGTGAAAAGGTTGTCCAGGAAGCTCTGGACAAAGCCAGAGAAGGC 3773

Qy     607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
           ||||||||:::|||:::         |||||||||||||||||||||||||||
Db    3774 CGCACCTGCATTGTGATCGCTCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTG 3833

Qy     627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
           |||:::   ||||||||:::||||||   ||||||||||||:::||||||      :::
Db    3834 GTGATTCAGAATGGCAAGGTCAAGGAGCATGGCACCCACCAGCAGCTGCTGGCACAGAAA 3893

Qy     647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
              |||||||||   :::|||:::   |||:::
Db    3894 GGCATCTATTTCTCCATGGTCAGTGTGCAGGCT 3926
```

```
RESULT 13
MUSMDRA
LOCUS       MUSMDRA                 4084 bp    mRNA    linear   ROD 27-APR-1993
DEFINITION  Mouse mdr gene encoding a multidrug resistance protein mRNA,
            complete cds.
ACCESSION   J03398
VERSION     J03398.1  GI:199109
KEYWORDS    multidrug resistance protein.
SOURCE      Mouse (strain BALB/c) pre-B-cell (cell line 702/3), cDNA to mRNA,
            clone MDR2.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4084)
  AUTHORS   Gros,P., Raymond,M., Bell,J. and Housman,D.
  TITLE     Cloning and characterization of a second member of the mouse mdr
            gene family
  JOURNAL   Mol. Cell. Biol. 8 (7), 2770-2778 (1988)
  MEDLINE   88302195
   PUBMED   3405218
REFERENCE   2  (sites)
  AUTHORS   Raymond,M. and Gros,P.
  TITLE     Mammalian multidrug-resistance gene: correlation of exon
            organization with structural domains and duplicaticn of an
            ancestral gene
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 86 (17), 6488-6492 (1989)
  MEDLINE   89367274
   PUBMED   2570420
REFERENCE   3  (sites)
  AUTHORS   Raymond,M. and Gros,P.
  TITLE     Cell-specific activity of cis-acting regulatory elements in the
            promoter of the mouse multidrug resistance gene mdr1
  JOURNAL   Mol. Cell. Biol. 10 (11), 6036-6040 (1990)
  MEDLINE   91042535
   PUBMED   2248681
COMMENT     Submitted in computer readable form by P.Gros 08-JUL-1988.
FEATURES             Location/Qualifiers
     source          1..4084
```

```
                  /organism="Mus musculus"
                  /strain="BALB/c"
                  /sub_species="domesticus"
                  /db_xref="taxon:10090"
                  /cell_line="702/3"
                  /cell_type="pre-B cell"
     CDS          194..4024
                  /codon_start=1
                  /product="multidrug resistance protein"
                  /protein_id="AAA39516.1"
                  /db_xref="GI:387428"
                  /translation="MDLEAARNGTARRLDGDFELGSISNQGREKKKKVNLIGLLTLFR
                  YSDWQDKLFMFLGTLMAIAHGSGLPLMMIVFGEMTDKFVDNTGNFSLPVNFSLSMLNP
                  GRILEEEMTRYAYYYSGLGGGVLVAAYIQVSFWTLAAGRQIKKIRQKFFHAILRQEMG
                  WFDIKGTTELNTRLTDDVSKISEGIGDKVGMFFQAIATFFAGFIVGFIRGWKLTLVIM
                  AISPILGLSTAVWAKILSTFSDKELAAYAKAGAVAEEAPGAIRTVIAFGGQNKELERY
                  QKHLENAKKIGIKKAISANISMGIAFLLIYASYALAFWYGSTLVISKEYTIGNAMTVF
                  FSILIGAFSVGQAAPCIDAFANARGAAYVIFDIIDNNPKIDSFSERGHKPDNIKGNLE
                  FSDVHFSYPSRANIKILKGLNLKVKSGQTVALVGNSGCGKSTTVQLLQRLYDPTEGKI
                  SIDGQDIRNFNVRCLREIIGVVSQEPVLFSTTIAENIRYGRGNVTMDEIEKAVKEANA
                  YDFIMKLPQKFDTLVGDRGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTESEA
                  EVQAALDKAREGRTTIVIAHRLSTIRNADVIAGFEDGVIVEQGSHSELMKKEGIYFRL
                  VNMQTAGSQILSEEFEVELSDEKAAGDVAPNGWKARIFRNSTKKSLKSPHQNRLDEET
                  NELDANVPPVSFLKVLKLNKTEWPYFVVGTVCAIANGALQPAFSIILSEMIAIFGPGD
                  DAVKQQKCNMFSLVFLGLGVLSFFTFFLQGFTFGKAGEILTTRLRSMAFKAMLRQDMS
                  WFDDHKNSTGALSTRLATDAAQVQGATGTKLALIAQNTANLGTGIIISFIYGWQLTLL
                  LLSVVPFIAVAGIVEMKMLAGNAKRDKKEMEAAGKIATEAIENIRTVVSLTQERKFES
                  MYVEKLHGPYRNSVRKAHIYGITFSISQAFMYFSYAGCFRFGSYLIVNGHMRFKDVIL
                  VFSAIVLGAVALGHASSFAPDYAKAKLSAAYLFSLFERQPLIDSYSGEGLWPDKFEGS
                  VTFNEVVFNYPTRANVPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERFYDPMAG
                  SVLLDGQEAKKLNVQWLRAQLGIVSQEPILFDCSIAENIAYGDNSRVVPHDEIVRAAK
                  EANIHPFIETLPQKYNTRVGDKGTQLSGGQKQRIAIARALIRQPRVLLLDEATSALDT
                  ESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVIENGKVKEHGTHQQLLAQKGI
                  YFSMVNIQAGTQNL"
BASE COUNT     1117 a    934 c   1075 g    958 t
ORIGIN     43 bp upstream of HinfI site.

Alignment Scores:
Pred. No.:              3.27e-145       Length:          4084
Score:                  1915.50         Matches:         366
Percent Similarity:     73.95%          Conservative:    128
Best Local Similarity:  54.79%          Mismatches:      163
Query Match:            57.64%          Indels:          11
DB:                     10              Gaps:            1

US-09-873-409-1 (1-659) x MUSMDRA (1-4084)

Qy       1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
             ::::::   |||:::|||:::|||:::||||||||    |||     |||:::|||:::
Db    2003 GTCATTGTGGAACAAGGAAGTCACAGTGAGCTGATGAAGAAGGAAGGGATCTACTTCAGA 2062

Qy      21 LeuValMetSerGln--------------------------------AspIleLysLys 29
           ||||||         |||                               :::::::::
Db    2063 CTCGTTAACATGCAGACAGCAGGAAGCCAGATCCTGTCAGAAGAATTTGAAGTTGAGCTA 2122

Qy      30 AlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPro 49
           :::||||||:::                                  |||
Db    2123 AGTGACGAAAAGGCTGCTGGAGATGTGGCCCCAAATGGCTGGAAAGCACGCATATTTAGG 2182

Qy      50 LeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGln 69
                 :::   ||||||:::||||||         ::::::    :::|||    |||
Db    2183 AATTCTACAAAGAAAAGTCTTAAAAGTCCACATCAGAATAGGCTGGATGAAGAAACCAAT 2242

Qy      70 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 89
                 :::   ::::::|||    ||||||    ||||||:::||||||||||||||||||
Db    2243 GAACTTGATGCAAACGTGCCACCAGTGTCTTTTCTGAAGGTCTTAAAACTGAATAAAACA 2302

Qy      90 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 109
           |||||||||:::    |||:::|||||||:::    ::::::    ||||||    :::    |||
Db    2303 GAGTGGCCCTACTTTGTGGTGGGAACAGTCTGTGCCATTGCCAATGGAGCCCTCCAGCCG 2362

Qy     110 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 129
              ||||||||||||    :::::::::|||    :::||||||       |||
Db    2363 GCTTTCTCCATCATCCTGTCTGAGATGATAGCTATCTTTGGCCCTGGGGATGACGCAGTG 2422

Qy     130 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 149
              :::             ::::::|||::::::|||:::   |||||||||:::   |||
Db    2423 AAGCAGCAAAAGTGTAACATGTTCTCCCTGGTCTTCTTGGGCCTAGGAGTCCTCTCCTTC 2482

Qy     150 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 169
              ::::::|||:::||||||      :::|||:::||||||||||||||||||   |||
Db    2483 TTTACTTTCTTCCTTCAGGGCTTCACGTTTGGGAAAGCTGGAGAGATCCTCACCACAAGG 2542

Qy     170 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLys 189
           ||||||   :::||||||||||||||||||   ||||||::::::||||||||||:::
Db    2543 CTCCGGTCCATGGCCTTTAAAGCGATGCTAAGGCAGGACATGAGCTGGTTTGATGATCAT 2602

Qy     190 GluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 209
           :::||||||||||||   |||:::|||   ||||||   |||   ||||||:::||||||
Db    2603 AAAAACAGTACTGGAGCACTTTCTACAAGACTCGCCACAGATGCTGCGCAAGTCCAAGGA 2662

Qy     210 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 229
           |||||||||:::::::::   ::::::   ||||||      |||:::|||     :::
Db    2663 GCCACGGGAACCAAGTTGGCTTTAATTGCACAGAACACAGCCAACCTTGGAACCGGTATT 2722

Qy     230 IleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVal 249
           ||||||||||||||||||||||||:::::::|||   |||:::||||||:::    |||
Db    2723 ATTATATCATTTATTTACGGTTGGCAACTGACACTTCTGCTGTTATCGGTTGTTCCATTC 2782

Qy     250 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 269
           :::||||||   |||:::|||          :::   |||   |||    :::||||||
Db    2783 ATTGCTGTAGCAGGAATTGTTGAAATGAAAATGTTGGCTGGCAATGCCAAGAGAGATAAA 2842

Qy     270 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIle 289
           :::|||::::::   ||||||||||||||||||||||||:::|||||||||||||||:::
Db    2843 AAGGAAATGGAAGCTGCTGGAAAGATTGCAACAGAGGCAATAGAAAATATTCGAACTGTT 2902

Qy     290 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGln 309
           |||||||||||:::|||:::   ||||||   ||||||   |||   |||
Db    2903 GTATCCTTGACCCAAGAAAGAAAATTTGAGTCAATGTATGTTGAAAAATTGCATGGACCT 2962

Qy     310 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla 329
           :::||||||:::   :::||||||   |||   |||      ::::::   |||   |||
Db    2963 TACAGGAATTCGGTGCGGAAGGCACACATCTACGGCATCACTTTTAGCATCTCCCAAGCA 3022

Qy     330 PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGly 349
           |||:::||||||:::||||||      ||||||||||||:::|||||||||      |||
Db    3023 TTCATGTATTTTTCTTATGCTGGCTGTTTTCGATTTGGTTCTTACCTAATTGTGAATGGA 3082

Qy     350 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAla 369
              |||      :::   :::   :::||||||:::||||||      ||||||:::|||
Db    3083 CATATGCGCTTCAAAGATGTCATTCTGGTCTTTTCTGCAATTGTGCTTGGCGCGGTGGCT 3142

Qy     370 IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHis 389
           :::|||            ||||||:::|||:::|||||||||      ||||||:::
Db    3143 CTAGGACACGCCAGCTCATTTGCTCCGGACTATGCAAAAGCCAAGCTGTCTGCAGCATAC 3202

Qy     390 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLys 409
           ||||||:::|||   |||::::::|||   |||||||||   |||   ||||||
Db    3203 TTGTTCAGCCTGTTTGAAAGACAACCTCTGATTGACAGCTACAGTGGAGAAGGGCTGTGG 3262

Qy     410 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 429
           ||||||      ||||||::::::   |||      |||   ||||||      |||
Db    3263 CCTGATAAGTTTGAAGGAAGCGTGACATTTAATGAAGTCGTGTTCAACTATCCCACCCGG 3322

Qy     430 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 449
              :::|||   :::|||:::||||||||||||   :::::::::|||:::|||:::|||
Db    3323 GCCAACGTGCCAGTGCTTCAGGGGCTGAGCCTTGAGGTGAAGAAGGGCCAGACGCTGGCC 3382

Qy     450 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr 469
              ||||||||||||||||||||||||||||   ||||||||||||:::|||   |||
Db    3383 CTGGTGGGCAGCAGTGGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAGCGCTTCTAT 3442

Qy     470 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 489
           ||||||:::   |||   ||||||   ||||||   :::||||||:::||||||||||||
Db    3443 GACCCCATGGCTGGATCAGTGCTCTTAGATGGTCAAGAAGCAAAGAAACTCAATGTCCAG 3502
```

```
Qy     490 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 509
           ||||||||:::|||:::   ||||||   |||||||||:::||||||:::|||||||||
Db    3503 TGGCTCCGAGCTCAACTGGGCATTGTGTCCCAGGAACCCATTCTCTTTGACTGCAGCATC 3562

Qy     510 AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu 529
           ||||||||||||||||||||||||||||||||||||||||||||||||    |||||||||
Db    3563 GCAGAGAACATCGCCTATGGAGACAACAGCCGGGTCGTGCCTCATGATGAGATTGTGAGG 3622

Qy     530 AlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThr 549
           ||||||     ||||||||||||   |||||||||   ||||||:::||||||||||||
Db    3623 GCAGCCAAGGAGGCCAACATCCACCCCTTCATCGAGACGCTGCCCCAAAAATATAACACA 3682

Qy     550 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAla 569
           :::||||||   ||||||   |||||||||||||||||||||||||:::|||||||||
Db    3683 AGAGTAGGAGACAAGGGGACGCAGCTCTCTGGGGGCCAGAAGCAGAGGATTGCCATCGCC 3742

Qy     570 ArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsp 589
           |||||||||:::::::::|||::::::||||||||||||||||||||||||||||||||
Db    3743 CGAGCCCTCATCAGACAGCCTCGGGTCCTACTGCTGGATGAAGCCACGTCAGCTCTGGAT 3802

Qy     590 AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys 609
              :::||||||||||||||||||   ||||||||||||||||||   ||||||||||||
Db    3803 ACTGAGAGTGAAAAGGTTGTCCAGGAAGCACTGGACAAAGCCAGGGAAGGCCGCACCTGC 3862

Qy     610 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 629
           :::|||:::   ||||||||||||   ||||||||||||||||||||||||||||:::
Db    3863 ATTGTGATCGCTCACCGCCTGTCCACCATCCAGAACGCGGACTTGATCGTGGTGATTGAG 3922

Qy     630 AsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyr 649
           |||||||||:::||||||   ||||||||||||:::||||||      :::   ||||||
Db    3923 AACGGCAAGGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCGCAGAAGGGCATCTAT 3982

Qy     650 PheLysLeuValAsnAlaGlnSer 657
           |||   :::||||||   |||:::
Db    3983 TTCTCAATGGTCAACATCCAGGCC 4006
```

```
RESULT 14
GGA9799
LOCUS       GGA9799                 4018 bp    mRNA    linear   VRT 03-MAR-1999
DEFINITION  Gallus gallus mRNA for ABC transporter protein.
ACCESSION   AJ009799
VERSION     AJ009799.1  GI:3355756
KEYWORDS    ABC transporter protein; cmdr1 gene; P-glycoprotein.
SOURCE      Gallus gallus.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 4018)
  AUTHORS   Edelmann,H.M.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JUL-1998) Edelmann H.M.L., Institute of Molecular
            Genetics, Biocenter and University of Vienna, Dr. Bohr-Gasse 9/2
            A-1030 Vienna, A-1030 Vienna, AUSTRIA
REFERENCE   2  (bases 1 to 4018)
  AUTHORS   Edelmann,H.M., Duchek,P., Rosenthal,F.E., Foger,N., Glackin,C.,
            Kane,S.E. and Kuchler,K.
  TITLE     Cmdr1, a chicken P-glycoprotein, confers multidrug resistance and
            interacts with estradiol
  JOURNAL   Biol. Chem. 380 (2), 231-241 (1999)
  MEDLINE   99209805
   PUBMED   10195430
FEATURES             Location/Qualifiers
     source          1..4018
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /tissue_type="intestine"
     gene            109..3975
                     /gene="cmdr1"
     CDS             109..3975
                     /gene="cmdr1"
                     /note="ABC transporter protein"
                     /codon_start=1
                     /product="P-glycoprotein"
                     /protein_id="CAA08835.1"
                     /db_xref="GI:3355757"
                     /db_xref="SPTREMBL:O93437"
                     /translation="MHSEDEKQKHTVDGNTYEIANSNQDQDPEDEKKGKKKKGKKPQM
                     VSPLALFRYSSCTDKLLMIFGSLLAIAHGTSLPIAMIIFGDMTDSFVTSGMTNITGNS
                     SGLNSSADVFNKLEEEMTRYAYYYSAIAAAVLVAAYIQTSFWTLAAGRQVKKIREKFF
                     HAIMRQEIGWFDVNDAGELNTRLIDDVSKINEGIGDKIGFLIQSETTFLTGFIVGFIR
                     GWKLTLVILAVSPVLGLSAALWAKILTAFTDKEQAAYAKAGAVAEEVLSAVRTVIAFG
                     GQEKEIKRYHKNLEDAKRIGIRKAITSNISMGAAFLLIYASYALAFWYGTTLILANEY
                     SIGNVLTVFFSVLIGAFSIGQTAPSIEAFANARGAAYAIFNIIDNEPEIDSYSDAGHK
                     PDHIKGNLEFQNVFFNYPSRPDVEILKGLNLKVNCGQTVALVGGSGCGKSTTVQLIQR
                     FYDPKEGTITIDGQDLKSLNVRYLREIIGVVNQEPVLFATTIAENIRYGREDVTMEEI
                     ERATKEANAYDFIMKLPKKFETVVGERGAQMSGGQKQRIAIARALVHNPKILLLDEAT
                     SALDTESESVVQAALDKAREGRTTVVVAHRLSTVRNADLIAVFESGVITEQGNHSQLI
                     EKKGIYYKLVNMQTIETEDPSSEKSENAVSVKRSGSQSNLDESLKKELRRGSTRRSMK
                     KPGEPNDTDEKGSSPDEELPPVSFLKLMKLNKNEWPYFVAGTFCAIVNGALQPAFSVI
                     FSEIIGIFSETDQKVLREKSNLYSLLFLALGIISFFTFFVQGFAFGKAGEILTMKLRF
                     MAFKAMLRQDMAWFDDPKNSTGALTTRLANDASQVKGATGVRLALIAQNIANLGTGII
                     ISLVYGWQLTLLLLAVVPIIAVAGMIEMKMLAGHAKKDKIELEAAGKIATEAIENIRT
                     VASLTREKRFELMYGEHLLVPYRNSVKKAHIFGFCFSLSQAMMFFTYAGCFRFGAYLV
                     VNGHIEYKTVFLVFSAVVFGAMALGQTSSFAPDYAKAKISAAHLFVLFNRVPPIDSYR
                     EDGEKPEKFGGNTRIKDVKFNYPNRPEVKILQGLNLAVEKGETLALVGSSGCGKSTVV
                     QLLERFYDPLSGEIVFDDIDAKTLNIQWLRSHIGIVSQEPILFDFTIAENIAYGDNSR
                     EVSHEEIISAAKAASIHSFIDSLPEKYNTRVGDKGTQLSGGQKQRIAIARALIRKPQI
                     LLLDEATSALDTESEKIVQEALDKAREGRTCIVIAHRLSTIQNADKIAVIQNGKVIEQ
                     GTHQQLLAEKGFYYSLVNVQSGSCNM"
     mat_peptide     112..3972
                     /gene="cmdr1"
                     /product="ABC transporter protein"
                     /note="P-glycoprotein"
BASE COUNT     1288 a    758 c    911 g   1061 t
ORIGIN

Alignment Scores:
Pred. No.:              7.4e-145        Length:         4018
Score:                  1911.00         Matches:        362
Percent Similarity:     72.88%          Conservative:   127
Best Local Similarity:  53.95%          Mismatches:     168
Query Match:            57.51%          Indels:         14
DB:                     5               Gaps:           1

US-09-873-409-1 (1-659) x GGA9799 (1-4018)
```

```
Qy       1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
           ::::::   |||:::|||   |||::::::|||:::   |||:::|||:::||||||
Db    1945 GTTATCACAGAACAAGGAAATCATTCTCAACTGATTGAAAAAAAGGGAATCTACTATAAA 2004

Qy      21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
           ||||||      |||   |||:::      |||         |||            :::
Db    2005 CTTGTTAACATGCAGACAATAGAAACTGAAGATCCATCATCAGAAAAAAGTGAAAATGCA 2064

Qy      41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
                    |||   :::         :::      :::|||:::|||   :::
Db    2065 GTCAGTGTCAAAAGAAGTGGATCTCAATCAAATCTGGATGAATCCCTAAAAAAAGAATTG 2124

Qy      61 ------------------------------------------IleAspLysAlaGluGlu 66
                                                         :::      :::|||
Db    2125 AGAAGAGGATCAACTAGGAGAAGCATGAAAAAGCCAGGAGAACCTAATGATACTGATGAG 2184

Qy      67 SerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeu 86
              |||   :::      ||||||   ||||||   ||||||:::::::::||||||
Db    2185 AAAGGTAGTTCACCTGATGAAGAATTACCTCCAGTTTCATTTCTGAAACTTATGAAGTTG 2244

Qy      87 AsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThr 106
           ||||||   |||||||||:::   |||   ||||||      :::::::::||||||
Db    2245 AACAAAAACGAATGGCCATACTTTGTAGCTGGAACTTTCTGTGCAATTGTCAATGGAGCT 2304

Qy     107 ValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAsp 126
           :::   |||   ||||||:::||||||::::::||||||   :::|||         |||
Db    2305 TTACAACCTGCATTTTCAGTCATATTCTCAGAAATTATAGGGATTTTTTCAGAAACTGAT 2364
```

```
Qy    127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db   2365 CAGAAAGTTTTGAGAGAAAAGAGCAACTTATATTCACTGCTGTTTTTAGCACTTGGGATC 2424

Qy    147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db   2425 ATTTCCTTTTTACTTTCTTTGTTCAGGGTTTTGCATTTGGCAAAGCTGGAGAAATTCTT 2484

Qy    167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
Db   2485 ACAATGAAGCTTCGATTCATGGCATTTAAAGCAATGCTGAGACAGGACATGGCCTGGTTT 2544

Qy    187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db   2545 GATGATCCTAAAAATAGCACTGGAGCACTGACAACAAGACTTGCTAATGATGCCTCACAA 2604

Qy    207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db   2605 GTAAAAGGAGCAACTGGTGTCAGATTGGCATTAATTGCCCAAAACATAGCTAACCTGGGA 2664

Qy    227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
Db   2665 ACTGGAATTATTATATCTTTAGTTTATGGCTGGCAGCTGACCCTGTTACTCTTAGCTGTT 2724

Qy    247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db   2725 GTACCAATCATTGCAGTGGCAGGAATGATTGAAATGAAGATGTTGGCTGGGCATGCAAAA 2784

Qy    267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db   2785 AAAGATAAAATAGAGCTGGAAGCTGCAGGAAAGATTGCCACAGAAGCCATAGAAAATATT 2844

Qy    287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db   2845 CGAACTGTTGCTAGTTTGACTCGGGAAAAAAGATTTGAGCTCATGTACGGAGAACATTTG 2904

Qy    307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db   2905 CTTGTACCATACAGAAATTCTGTGAAGAAGGCGCACATATTTGGGTTTTGTTTTTCCCTT 2964

Qy    327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
Db   2965 TCACAAGCAATGATGTTTTTTACCTATGCTGGCTGTTTCCGCTTTGGAGCCTATCTTGTG 3024

Qy    347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db   3025 GTTAATGGACACATAGAGTACAAAACTGTGTTTTTAGTGTTTTCAGCTGTTGTATTTGGT 3084

Qy    367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db   3085 GCTATGGCATTAGGACAAACCAGCTCTTTTGCACCAGACTATGCCAAAGCCAAGATATCA 3144

Qy    387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Db   3145 GCAGCCCACTTGTTTGTGCTGTTTAATAGAGTACCTCCAATAGACAGTTACCGCGAGGAT 3204

Qy    407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
Db   3205 GGAGAGAAACCTGAAAAATTTGGGGGAAACACAAGGATCAAAGATGTGAAATTTAACTAT 3264

Qy    427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db   3265 CCAAACCGACCAGAGGTCAAAATCCTCCAAGGCTTGAATCTAGCAGTAGAAAAAGGAGAA 3324

Qy    447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
Db   3325 ACGTTGGCCCTTGTTGGTAGCAGTGGATGTGGAAAGAGCACTGTTGTTCAGCTACTTGAG 3384

Qy    467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db   3385 AGATTTTATGATCCACTCAGTGGAGAAATTGTGTTTGATGACATTGATGCAAAGACACTA 3444

Qy    487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db   3445 AATATCCAGTGGCTGAGATCTCACATCGGTATCGTCTCACAAGAGCCAATCCTGTTCGAC 3504

Qy    507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db   3505 TTCACCATCGCTGAAAACATCGCGTACGGGGACAACAGTCGGGAGGTGTCCCATGAGGAA 3564

Qy    527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db   3565 ATCATTAGTGCAGCGAAAGCAGCCAGTATTCATTCCTTCATTGACTCTCTGCCAGAAAAA 3624

Qy    547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Db   3625 TACAATACTCGTGTAGGAGACAAGGGAACACAGCTTTCTGGTGGTCAGAAACAACGTATT 3684

Qy    567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Db   3685 GCTATTGCTCGAGCTCTTATACGTAAGCCCCAGATTCTGCTACTGGATGAAGCTACATCT 3744

Qy    587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
Db   3745 GCCTTGGACACAGAAAGCGAAAAGATTGTCCAGGAAGCACTGGATAAAGCCAGAGAAGGT 3804

Qy    607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db   3805 CGCACCTGCATCGTGATAGCTCACCGCCTGTCCACCATCCAGAATGCTGACAAGATTGCT 3864

Qy    627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
Db   3865 GTGATCCAGAACGGCAAGGTCATAGAGCAAGGGACTCACCAGCAGCTCCTGGCTGAAAAA 3924

Qy    647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db   3925 GGCTTCTACTACTCGCTGGTCAATGTTCAAAGT 3957
```

```
RESULT 15
OAU78609
LOCUS       OAU78609                3858 bp    mRNA    linear   MAM 11-OCT-2001
DEFINITION  Ovis aries multidrug resistance protein-1 (mdr1) mRNA, complete
            cds.
ACCESSION   U78609
VERSION     U78609.1  GI:2149086
KEYWORDS    .
SOURCE      Ovis aries.
  ORGANISM  Ovis aries
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Caprinae; Ovis.
REFERENCE   1  (bases 1 to 3858)
  AUTHORS   Longley,M., Phua,S.H., van Stijn,T.C. and Crawford,A.M.
  TITLE     Isolation and mapping of the first ruminant multidrug resistance
            genes
  JOURNAL   Anim. Genet. 30 (3), 207-210 (1999)
  MEDLINE   99371931
   PUBMED   10442984
REFERENCE   2  (bases 1 to 3858)
  AUTHORS   Longley,M. and Crawford,A.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-NOV-1996) Biochemistry Department, University of
            Otago, Dunedin, New Zealand
FEATURES             Location/Qualifiers
     source          1..3858
                     /organism="Ovis aries"
                     /db_xref="taxon:9940"
     gene            1..3858
                     /gene="mdr1"
     CDS             1..3858
                     /gene="mdr1"
                     /codon_start=1
                     /product="multidrug resistance protein-1"
                     /protein_id="AAB58489.1"
                     /db_xref="GI:2149087"
                     /translation="MDLEGDRNGRAAGGNFLKRDKKRFFSKKDEKKEKRPTVSTFTMF
```

```
                    RYSNWLDRLCMVLGTLAAIIHGAGLPLMMLVFGDMTDSFAGAGNLGNITLSNISNTST
                    IDRTEYGKKLEKEMTTYAYYYSGIGAGVLIAAYIQVSFWCLAAGRQVHRIRKQFFHAI
                    MQQEIGWFDVHDVGKLNTRLTNDVSKINEGIGDKIGMFFQAMATFFTGFIIGFTTGWN
                    LTLVILAIIPVLGLSAAIWAKILSSFTDKKLLAYAKAGAVAEEVLAAIKTVIAFGGQK
                    KELERYNKNLEEAKRIGIKKAITANISMGAAFLLIYASYALAFWYGTSLVLSREYSIG
                    QVLTVFFSVLIGAFSIGQASPNIEAFANARGAAYEVFKIIDNKPSIDSYSNTGHKPDN
                    IKGNLEFRNVHFHYPSRNEVKILKGLNLKVGSGQTVALVGNSGCGKSTTVQLMQRLYD
                    PTEGMVSIDGQDIRTINVRYLREIIGVVSQEPVLFATTIAENIRYGREDVTMDEIQKA
                    VKEANAYDFIMKLPNKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSAL
                    GTESEAVVQAALDKARKGRTTIVIAHRLSTVRNADIIAGLDDGVIVEEGSHDELMGKR
                    GIYFKLVTMQTKGNELELENTPGESLSKTDDLYMSSQDSRSSLIRRKSTRRSIRGSQS
                    QDRKLSTEETLDESVPPVSFWRILKLNITEWPYFVVGVFCAIINGALQPAFSVIFSRI
                    IGIFTRNDNDETKRQNSNLFSLLFLILGIISFITFFLQGFTFGKAGEILTRRLRYLVF
                    RSMLRQDVSWFDDPKNTTGALTTRLANDAAQVKGAVGSRLAVITQNIANLGTGIIISL
                    IYGWQLTLLLLAIVPIIAVAGVIEMKMLSGQALKDKKELEGAGKIATEAIENFRTVVS
                    LTREERFEYMYAQSLQVPYRNSLRRAHVFGITFSITQAMMYFSYAGCFRFGAYLVAQG
                    IMEFQDVLLVFSAVVFGAMAVGQVSTFAPDYAKAKVSAAHVINIIEKIPLIDSYSTEG
                    LKPSTVEGSVAFNDVVFNYPTRPDVPVLRGLSLEVKKGQTLALVGSSGCGKSTVVQLL
                    ERFYDPLAGTVFIDGKEVKQLNVQWLRAHMGIVSQEPILFDCSIGENIAYGDNSRVVS
                    QEEIEHAAKEANIHSFIEMLPDKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLL
                    DEATSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGRIKEHGTH
                    QQLLAQKGIYFTMVSVQAGTKRQ"
BASE COUNT     1072 a    853 c    1002 g     931 t
ORIGIN

Alignment Scores:
Pred. No.:               2e-143          Length:         3858
Score:                   1893.00         Matches:        364
Percent Similarity:      73.85%          Conservative:   133
Best Local Similarity:   54.09%          Mismatches:     158
Query Match:             56.97%          Indels:         18
DB:                      4               Gaps:           4

US-09-873-409-1 (1-659) x OAU78609 (1-3858)

Qy       1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db    1828 GTCATCGTGGAGGAAGGCAGTCATGATGAGCTCATGGGGAAAAGAGGCATTTACTTCAAA 1887

Qy      21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
Db    1888 CTTGTCACAATGCAG---ACAAAAGGAAATGAACTTGAATTGGAAAATACCCCTGGTGAA 1944

Qy      41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
Db    1945 TCCCTAAGTAAAACTGATGACTTGTACATGTCCTCCCAAGATTCA---AGATCCAGTCTA 2001

Qy      61 Ile---------------------------------------------AspLysAlaGlu 65
Db    2002 ATCAGAAGAAAATCAACTCGCAGGAGTATCCGTGGATCCCAAAGCCAGGACAGAAAGCTC 2061

Qy      66 GluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLys 85
Db    2062 AGTACAGAAGAAACTTTGGATGAAAGTGTACCTCCCGTTTCCTTTTGGAGGATTCTGAAG 2121

Qy      86 LeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGly 105
Db    2122 CTGAATATAACTGAATGGCCTTATTTGTGGTTGGTGTATTTGCGCCATTATAAATGGA 2181

Qy     106 ThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsn 125
Db    2182 GCTCTGCAACCAGCATTCTCAGTCATATTTTCAAGGATTATAGGAATCTTCACCAGAAAT 2241

Qy     126 AspLys---ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeu 144
Db    2242 GACAATGACGAAACCAAACGACAGAACAGTAACTTGTTTTCACTGTTGTTTCTAATCCTT 2301

Qy     145 GlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGlu 164
Db    2302 GGAATTATTTCTTTTATTACATTTTTCCTTCAGGGCTTCACGTTCGGCAAAGCAGGCGAG 2361

Qy     165 IleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAla 184
Db    2362 ATCCTCACCAGGCGGCTGCGGTACCTGGTTTTCAGGTCGATGCTGAGACAGGATGTGAGC 2421

Qy     185 TrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIle 204
Db    2422 TGGTTCGATGACCCTAAAAACACCACGGGAGCATTGACGACCAGGCTTGCCAATGACGCC 2481

Qy     205 AlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsn 224
Db    2482 GCTCAAGTTAAAGGGGCTGTAGGTTCCAGGCTTGCTGTAATTACCCAGAATATTGCAAAT 2541

Qy     225 MetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeu 244
Db    2542 CTTGGCACAGGCATTATCATATCCCTCATCTATGGCTGGCAGTTAACACTTCTCCTGTTA 2601

Qy     245 SerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPhe 264
Db    2602 GCAATTGTACCCATCATTGCAGTAGCAGGAGTTATTGAAATGAAAATGCTGTCTGGACAA 2661

Qy     265 AlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGlu 284
Db    2662 GCCCTGAAAGATAAGAAAGAACTGGAAGGTGCTGGGAAGATTGCCACTGAAGCAATAGAA 2721

Qy     285 AsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGlu 304
Db    2722 AACTTCCGGACGGTTGTTTCTTTGACTCGGGAGGAGAGGTTTGAATATATGTATGCCCAG 2781

Qy     305 MetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyr 324
Db    2782 AGTTTGCAGGTACCATACAGAAACTCCTTGAGGAGAGCGCATGTCTTTGGAATTACATTT 2841

Qy     325 AlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyr 344
Db    2842 TCCATCACTCAGGCAATGATGTATTTCTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTAC 2901

Qy     345 LeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAla 364
Db    2902 TTGGTGGCCCAAGGCATCATGGAGTTTCAGGATGTTCTCTTAGTGTTCTCAGCTGTTGTC 2961

Qy     365 TyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLys 384
Db    2962 TTTGGTGCCATGGCTGTGGGACAGGTCAGTACATTTGCTCCTGACTATGCCAAGGCCAAA 3021

Qy     385 SerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSer 404
Db    3022 GTGTCAGCAGCCCACGTCATCAATATCATTGAAAAAATTCCTCTGATTGACAGCTACAGC 3081

Qy     405 GlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhe 424
Db    3082 ACGGAAGGCCTAAAGCCGAGTACAGTGGAAGGAAGTGTGGCATTTAATGACGTCGTGTTC 3141

Qy     425 PheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArg 444
Db    3142 AACTACCCCACTCGACCAGACGTCCCAGTGCTTCGGGGGCTGAGCCTGGAGGTGAAGAAG 3201

Qy     445 GlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeu 464
Db    3202 GGCCAGACGCTGGCGCTGGTGGGCAGTAGTGGCTGTGGGAAGAGCACGGTCGTCCAGCTC 3261

Qy     465 LeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLys 484
Db    3262 CTGGAGCGGTTCTATGACCCCTTGGCTGGCACAGTGTTTATTGACGGCAAGGAAGTAAAG 3321

Qy     485 GluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeu 504
Db    3322 CAGCTGAATGTCCAGTGGCTGCGGGCACACATGGGCATCGTGTCCCAGGAGCCCATTCTG 3381

Qy     505 PheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeu 524
Db    3382 TTTGACTGCAGCATCGGCGAGAATATCGCCTACGGAGACAACAGCCGGGTCGTGTCCCAG 3441

Qy     525 AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuPro 544
Db    3442 GAGGAGATTGAGCACGCGGCCAAGGAGGCCAACATCCACTCCTTCATCGAGATGCTGCCG 3501
```

```
Qy     545 GluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGln 564
           :::||||||||||||:::||||||   ||||||   ||||||||||||||||||||||||
Db    3502 GACAAATACAACACCAGAGTAGGGGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAG 3561

Qy     565 ArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAla 584
           |||:::||||||||||||||||||:::::::::|||   |||||||||||||||||||||
Db    3562 CGCATTGCCATAGCTCGCGCCCTTGTTAGACAGCCTCACATCTTGCTTCTGGATGAAGCT 3621

Qy     585 ThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArg 604
           |||||||||||||||   :::||||||||||||||||||   ||||||||||||||||||
Db    3622 ACCTCGGCTCTGGACACGGAAAGTGAAAAGGTGGTCCAAGAAGCCCTGGACAAAGCCCGA 3681

Qy     605 ThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeu 624
              ||||||||||||:::|||:::   ||||||||||||   ||||||||||||||||||
Db    3682 GAGGGACGCACCTGCATCGTGATCGCTCACCGGCTGTCCACCATCCAGAACGCAGACTTG 3741

Qy     625 IleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArg 644
           |||||||||      ||||||:::|||||||||   ||||||||||||:::||||||
Db    3742 ATAGTGGTGTTTCAGAATGGCAGAATCAAGGAGCATGGCACACACCAGCAGCTGCTGGCG 3801

Qy     645 AsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 657
              :::   |||||||||   :::|||:::   |||:::
Db    3802 CAGAAAGGCATCTATTTCACCATGGTCAGTGTCCAGGCT 3840
```

Search completed: March 31, 2003, 02:07:24
Job time : 8998.97 secs

```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 30, 2003, 01:20:01 ; Search time 644.062 Seconds
                                           (without alignments)
                                           2304.231 Million cell updates/sec

Title:          US-09-873-409-1
Perfect score:  3323
Sequence:       1 MLAEKGAHAELMAKRGLYYS..........QELLRNRDIYFKLVNAQSVQ 659

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters:      4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115416_19216/app_query.fasta_1.7544
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_4247_@runat_27032003_115416_19216 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      N_Geneseq_101002:*
                1:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
                2:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
                3:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
                4:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
                5:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
                6:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
                7:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
                8:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
                9:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
                10:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
                11:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
                12:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
                13:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
                14:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
                15:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
                16:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
                17:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
                18:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
                19:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
                20:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
                21:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
                22:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
                23:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
                24:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3163	95.2	3699	24	ABK83223	Human transporter
2	1924.5	57.9	3924	21	AAZ94742	Human ATP binding
3	1924.5	57.9	3924	21	AAZ88974	Human MDR-3 DNA.
4	1924.5	57.9	3924	24	ABN95801	Gene #2299 used to
5	1882.5	56.7	3912	24	ABK63653	Rat sequence diffe
6	1882	56.6	4369	21	AAZ52047	Rat multidrug resi
7	1882	56.6	4425	21	AAZ52048	Rat multidrug resi
8	1873	56.4	4186	22	AAF86127	Cynomologous monke
9	1873	56.4	4195	22	AAF86128	Cynomologous monke
10	1872	56.3	2726	15	AAQ70907	Multidrug-resistan
11	1872	56.3	2726	15	AAQ70916	Multidrug-resistan
12	1872	56.3	2726	18	AAT43322	Multidrug resistan
13	1872	56.3	3840	24	ABL91687	Human polynucleoti
14	1872	56.3	3860	21	AAZ49332	Human wild-type mu
15	1872	56.3	3860	21	AAZ49333	Human G185V mutant
16	1872	56.3	3860	24	ABA94365	Human BCRP DNA rel
17	1872	56.3	3860	24	ABA94366	Human BCRP DNA rel
18	1872	56.3	3988	21	AAZ88973	Human MDR-1 DNA.
19	1872	56.3	4264	19	AAV66533	Mutated human P-gl
20	1872	56.3	4264	19	AAV66534	Mutated human P-gl
21	1872	56.3	4349	22	AAH57442	Human intestine ce
22	1872	56.3	4646	21	AAZ94738	Human ATP binding
23	1872	56.3	4646	24	AAD38994	Human mdr1 gene.
24	1872	56.3	4646	24	ABL68592	Kidney cancer rela
25	1872	56.3	4646	24	ABL68880	Kidney cancer rela
26	1872	56.3	4669	8	AAN70752	Sequence of human
27	1872	56.3	6505	17	AAT13394	Hybrid vector pSF-
28	1872	56.3	8630	21	AAZ24041	Retroviral M4 mdr-
29	1872	56.3	8630	21	AAZ24042	Retroviral vector
30	1871	56.3	4788	21	AAZ49335	Murine multidrug r
31	1871	56.3	4788	24	ABA94368	Mouse BCRP DNA rel
32	1870	56.3	4279	22	AAD03488	Dog P-glycoprotein
33	1870	56.3	4279	22	AAD03504	Dog P-glycoprotein
34	1870	56.3	4279	22	AAD03505	Dog P-glycoprotein
35	1870	56.3	4317	22	AAD03489	Dog P-glycoprotein
36	1869	56.2	4279	22	AAD03506	Dog P-glycoprotein
37	1869	56.2	4378	11	AAQ04522	Multidrug Resistan
38	1869	56.2	4669	14	AAQ52726	Sequence of human
39	1861	56.0	4646	15	AAQ72872	Human multidrug re
40	1855	55.8	4669	19	AAV32645	Human P glycoprote
41	1855	55.8	4669	24	ABK52041	cDNA encoding huma
42	1849.5	55.7	4233	21	AAZ90198	Rat mdr1b2 (multis
43	1849.5	55.7	4233	22	AAF27498	Rat mdr1b2 multidr
44	1844.5	55.5	4189	21	AAZ49334	Murine multidrug r
45	1844.5	55.5	4189	24	ABA94367	Mouse BCRP DNA rel

ALIGNMENTS

```
RESULT 1
ABK83223
ID   ABK83223 standard; cDNA; 3699 BP.
XX
AC   ABK83223;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human transporter and ion channel, TRICH14, Incyte ID 7472030CB1, cDNA.
XX
KW   Human; ss; gene; transporter and ion channel; TRICH; transport disorder;
KW   neurological disorder; muscle disorder; immunological disorder; cancer;
KW   scleroderma; systemic lupus erythematosus; allergy; leukaemia;
KW   cell proliferative disorder; cervical cancer; breast cancer;
KW   neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;
KW   myotonic dystrophy; catatonia; endocrine disorder; diabetes;
KW   Grave's disease; gastrointestinal disorder; Crohn's disease;
KW   renal disorder; Good pasture's syndrome; viral infection; cirrhosis;
```

```
KW   bacterial infection; fungal infection; parasitic infection;
KW   protozoal infection; helminthic infection; cardiovascular disorder;
KW   atherosclerosis; hepatic disease.
XX
OS   Homo sapiens.
XX
PN   WO200240541-A2.
XX
PD   23-MAY-2002.
XX
PF   25-OCT-2001; 2001WO-US46055.
XX
PR   27-OCT-2000; 2000US-243989P.
PR   03-NOV-2000; 2000US-245904P.
PR   09-NOV-2000; 2000US-247673P.
PR   17-NOV-2000; 2000US-249661P.
PR   20-NOV-2000; 2000US-252232P.
PR   01-DEC-2000; 2000US-250790P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Tang YT,  Yue H,  Nguyen DB,  Hafalia AJA,  Elliott VS,  Lu Y;
PI   Walia NK,  Yao MG,  Baughn MR,  Gandhi AR,  Ding L,  Sanjanwala M;
PI   Ramkumar J,  Arvizu C,  Gietzen KJ,  Lal PG,  Azimzai Y,  Khan FA;
PI   Thangavelu K,  Thornton M,  Lu DAM,  Tribouley CM,  Warren BA;
PI   Ison CH,  Das D,  Raumann BE,  Policky JL,  Kearney L;
XX
DR   WPI; 2002-463570/49.
DR   P-PSDB; ABG61544.
XX
PT   New transporters and ion channels (TRICH) polypeptides, useful for
PT   diagnosing, preventing, and treating disorders associated with an
PT   abnormal expression or activity of TRICH, e.g. immunological, muscular
PT   or renal disorders  -
XX
PS   Claim 5; Page 172; 178pp; English.
XX
CC   The invention relates to human transporters and ion channels (TRICH)
CC   polypeptides, a naturally occurring amino acid sequence 90 % identical to
CC   TRICH, a biologically active fragment of TRICH or an immunogenic fragment
CC   of TRICH. Also included are an isolated polynucleotide encoding TRICH,
CC   a recombinant polynucleotide comprising a promoter sequence operably
CC   linked to the TRICH polynucleotide, a cell transformed with the
CC   recombinant polynucleotide, a transgenic organism comprising the
CC   recombinant polynucleotide, an isolated antibody that binds specifically
CC   to TRICH, and screening for compounds which bind to TRICH, modulate
CC   TRICH, modulate TRICH expression or are ant/agonists of TRICH.
CC   The polypeptides are useful for diagnosing, treating, and
CC   preventing transport, neurological, muscle, immunological disorders
CC   (e.g. scleroderma, systemic lupus erythematosus, allergies), cell
CC   proliferative disorders such as cancers (e.g. leukaemia, cervical or
CC   breast cancers), neurodegenerative disorders (e.g. Parkinson's disease,
CC   Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy,
CC   catatonia), endocrine disorders (e.g. diabetes, Grave's disease),
CC   gastrointestinal disorders (e.g. Crohn's disease), renal disorders
CC   (e.g. Good pasture's syndrome), viral, bacterial, fungal, parasitic,
CC   protozoal and helminthic infections, cardiovascular disorders (e.g.
CC   atherosclerosis), or hepatic diseases (e.g. cirrhosis) and many
CC   other diseases and disorders detailed in the specification. They can also
CC   be used in assessing the effects of exogenous compounds on the
CC   expression of nucleic acid and amino acid sequences of transporters and
CC   ion channels. TRICH or its fragments may also be used in screening for
CC   compounds that specifically bind to and modulate the activity of TRICH.
CC   The polynucleotides can be used to create knock-in humanised animals or
CC   transgenic animals to model human disease. The present sequence
CC   encodes a TRICH protein.
XX
SQ   Sequence 3699 BP; 1116 A; 707 C; 860 G; 1016 T; 0 other;

Alignment Scores:
Pred. No.:            2.7e-314     Length:        3699
Score:                3163.00      Matches:       632
Percent Similarity:   96.05%       Conservative:  1
```

```
Best Local Similarity:  95.90%      Mismatches:   0
Query Match:            95.19%      Indels:       26
DB:                     24          Gaps:         1

US-09-873-409-1 (1-659) x ABK83223 (1-3699)

Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1798 ATGCTGGCGGAGAAAGGAGCGCATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCA 1857

Qy     21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1858 CTTGTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCT 1917

Qy     41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1918 ACTGAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTC 1977

Qy     61 IleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1978 ATTGACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTA 2037

Qy     81 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2038 TTAAAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCT 2097

Qy    101 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2098 TCTGTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACC 2157

Qy    121 MetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2158 ATGTTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATA 2217

Qy    141 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 160
          ||||||||||||||||||||||||||||||||||||||||||
Db   2218 TTCGTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATG------------------ 2259

Qy    161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180

Db   2259 ------------------------------------------------------------ 2259

Qy    181 GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2260 CAGGATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTA 2319

Qy    201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2320 GCCATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAA 2379

Qy    221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2380 AATGCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACA 2439

Qy    241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2440 TTCCTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCA 2499

Qy    261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2500 ATGACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACT 2559

Qy    281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2560 GAAGCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAA 2619

Qy    301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2620 ATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATT 2679

Qy    321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2680 GGAAGCTGTTATGCATTCAGCCATGCCTTTATATATTTGCCTATGCGGCAGGGTTTCGA 2739
```

```
Qy     341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2740 TTTGGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTT 2799

Qy     361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380
           ||||||||||||||||||||||||||||||||||:::|||||||||||||||||||||||
Db    2800 ACTGCAATTGCATATGGAGCTATGGCCATCGGAGAAACGCTCGTTTTGGCTCCTGAATAT 2859

Qy     381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2860 TCCAAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATA 2919

Qy     401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2920 GACAGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGA 2979

Qy     421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2980 GAAGTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTC 3039

Qy     441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3040 AGTATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACT 3099

Qy     461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3100 TCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGT 3159

Qy     481 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3160 GTGGATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAA 3219

Qy     501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3220 GAGCCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGT 3279

Qy     521 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3280 GTGGTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATT 3339

Qy     541 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3340 GAAGGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGC 3399

Qy     561 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3400 GGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTG 3459

Qy     581 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3460 TTGGATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTT 3519

Qy     601 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3520 GATAAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAG 3579

Qy     621 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3580 AACGCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAA 3639

Qy     641 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 659
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3640 GAGCTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3696
```

RESULT 2
AAZ94742

```
ID   AAZ94742 standard; cDNA; 3924 BP.
XX
AC   AAZ94742;
XX
DT   01-AUG-2000  (first entry)
XX
DE   Human ATP binding cassette ABCB4 (MDR3) cDNA.
XX
KW   ABCB4; ATP binding cassette; human; cholesterol; lipid disorder;
KW   atherosclerosis; lipid disorder; dyslipidemia; psoriasis;
KW   lupus erythematosus; diagnosis; gene therapy; MDR3;
KW   multidrug resistance; chromosome 7q21; ss.
XX
OS   Homo sapiens.
XX
PN   WO200018912-A2.
XX
PD   06-APR-2000.
XX
PF   21-SEP-1999;   99WO-EP06991.
XX
PR   25-SEP-1998;   98US-0101706.
XX
PA   (FARB ) BAYER AG.
XX
PI   Schmitz G,  Klucken J;
XX
DR   WPI; 2000-293151/25.
XX
PT   Adenosine triphosphate binding proteins useful for identifying agents
PT   for treating atherosclerosis and other inflammatory disorders -
XX
PS   Claim 9; Page 117-119; 154pp; English.
XX
CC   The present sequence is that of human ATP binding cassette
CC   subfamily B protein ABCB4 cDNA.  The cDNA was identified using a
CC   differential display method in which monocytes from peripheral
CC   blood were subjected to macrophage differentiation and cholesterol
CC   loading with acetylated low density lipoproteins and subsequent
CC   deloading with high density lipoprotein (HDL3) to identify
CC   cholesterol sensitive genes.  The gene maps to chromosome 7q21
CC   and is also termed MDR3 (multidrug resistance).  The invention
CC   provides cholesterol-sensitive ABC genes (see AAZ94734-63).  These
CC   genes, and polypeptides encoded by them, can be used for diagnostic
CC   and therapeutic applications, and for biochemical or cell-based
CC   assays to screen for pharmacologically active modulator compounds
CC   useful for the treatment of lipid disorders, atherosclerosis or
CC   other inflammatory diseases such as psoriasis and lupus
CC   erythematosus.
XX
SQ   Sequence 3924 BP; 1145 A; 790 C; 977 G; 1012 T; 0 other;

Alignment Scores:
Pred. No.:              5.43e-187       Length:          3924
Score:                  1924.50         Matches:         372
Percent Similarity:     74.10%          Conservative:    123
Best Local Similarity:  55.69%          Mismatches:      162
Query Match:            57.91%          Indels:          11
DB:                     21              Gaps:            3

US-09-873-409-1 (1-659) x AAZ94742 (1-3924)

Qy       1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
           ::::::   |||:::|||:::|||:::|||||||||   |||   |||:::|||:::
Db    1851 GTAATTGTGGAGCAAGGAAGCCACAGCGAACTGATGAAGAAGGAAGGGGTGTACTTCAAA 1910

Qy      21 LeuVal---------------------------MetSerGlnAspIleLysLysAlaAsp 31
           ||||||                            |||::::::   :::      |||
Db    1911 CTTGTCAACATGCAGACATCAGGAAGCCAGATCCAGTCAGAAGAATTTGAACTAAATGAT 1970

Qy      32 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 51
           |||:::         :::                  |||:::
Db    1971 GAAAAGGCTGCCACTAGAATGGCCCCAAATGGCTGGAAATCTCGCCTATTTAGGCATTCT 2030

Qy      52 SerValLysSerIleLys---SerAspPheIleAspLysAla---GluGluSerThrGln 69
           :::   |||::::::|||   |||            |||:::   :::      |||
Db    2031 ACTCAGAAAAACCTTAAAAATTCACAAATGTGTCAGAAGAGCCTTGATGTGGAAACCGAT 2090

Qy      70 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 89
```

```
Db   2091 GGACTTGAAGCAAATGTGCCACCAGTGTCCTTTCTGAAGGTCCTGAAACTGAATAAAACA 2150

Qy     90 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 109
Db   2151 GAATGGCCCTACTTTGTCGTGGGAACAGTATGTGCCATTGCCAATGGGGGGCTTCAGCCG 2210

Qy    110 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 129
Db   2211 GCATTTTCAGTCATATTCTCAGAGATCATAGCGATTTTTGGACCAGGCGATGATGCAGTG 2270

Qy    130 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 149
Db   2271 AAGCAGCAGAAGTGCAACATATTCTCTTTGATTTTCTTATTTCTGGGAATTATTTCTTTT 2330

Qy    150 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 169
Db   2331 TTTACTTTCTTCCTTCAGGGTTTCACGTTTGGGAAAGCTGGCGAGATCCTCACCAGAAGA 2390

Qy    170 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLys 189
Db   2391 CTGCGGTCAATGGCTTTTAAAGCAATGCTAAGACAGGACATGAGCTGGTTTGATGACCAT 2450

Qy    190 GluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 209
Db   2451 AAAAACAGTACTGGTGCACTTTCTACAAGACTTGCCACAGATGCTGCCCAAGTCCAAGGA 2510

Qy    210 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 229
Db   2511 GCCACAGGAACCAGGTTGGCTTTAATTGCACAGAATATAGCTAACCTTGGAACTGGTATT 2570

Qy    230 IleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVal 249
Db   2571 ATCATATCATTTATCTACGGTTGGCAGTTAACCCTATTGCTATTAGCAGTTGTTCCAATT 2630

Qy    250 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 269
Db   2631 ATTGCTGTGTCAGGAATTGTTGAAATGAAATTGTTGGCTGGAAATGCCAAAAGAGATAAA 2690

Qy    270 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIle 289
Db   2691 AAAGAACTGGAAGCTGCTGGAAAGATTGCAACAGAGGCAATAGAAAATATTAGGACAGTT 2750

Qy    290 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGln 309
Db   2751 GTGTCTTTGACCCAGGAAAGAAAATTTGAATCAATGTATGTTGAAAAATTGTATGGACCT 2810

Qy    310 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla 329
Db   2811 TACAGGAATTCTGTGCAGAAGGCACACATCTATGGAATTACTTTTAGTATCTCACAAGCA 2870

Qy    330 PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGly 349
Db   2871 TTTATGTATTTTTCCTATGCCGGTTGTTTTCGATTTGGTGCATATCTCATTGTGAATGGA 2930

Qy    350 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAla 369
Db   2931 CATATGCGCTTCAGAGATGTTATTCTGGTGTTTTCTGCAATTGTATTTGGTGCAGTGGCT 2990

Qy    370 IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHis 389
Db   2991 CTAGGACATGCCAGTTCATTTGCTCCAGACTATGCTAAAGCTAAGCTGTCTGCAGCCCAC 3050

Qy    390 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLys 409
Db   3051 TTATTCATGCTGTTTGAAAGACAACCTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAG 3110

Qy    410 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 429
Db   3111 CCTGATAAATTTGAAGGAAATATAACATTTAATGAAGTCGTGTTCAACTATCCCACCCGA 3170

Qy    430 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 449
Db   3171 GCAAACGTGCCAGTGCTTCAGGGGCTGAGCCTGGAGGTGAAGAAAGGCCAGACACTAGCC 3230

Qy    450 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr 469
Db   3231 CTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACGGTGGTCCAGCTCCTGGAGCGGTTCTAC 3290

Qy    470 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 489
Db   3291 GACCCCTTGGCGGGGACAGTGCTTCTCGATGGTCAAGAAGCAAAGAAACTCAATGTCCAG 3350

Qy    490 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 509
Db   3351 TGGCTCAGAGCTCAACTCGGAATCGTGTCTCAGGAGCCTATCCTATTTGACTGCAGCATT 3410

Qy    510 AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu 529
Db   3411 GCCGAGAATATTGCCTATGGAGACAACAGCCGGGTTGTATCACAGGATGAAATTGTGAGT 3470

Qy    530 AlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThr 549
Db   3471 GCAGCCAAAGCTGCCAACATACATCCTTTCATCGAGACGTTACCCCACAAATATGAAACA 3530

Qy    550 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAla 569
Db   3531 AGAGTGGGAGATAAGGGGACTCAGCTCTCAGGAGGTCAAAAACAGAGGATTGCTATTGCC 3590

Qy    570 ArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsp 589
Db   3591 CGAGCCCTCATCAGACAACCTCAAATCCTCCTGTTGGATGAAGCTACATCAGCTCTGGAT 3650

Qy    590 AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys 609
Db   3651 ACTGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGC 3710

Qy    610 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 629
Db   3711 ATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAG 3770

Qy    630 AsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyr 649
Db   3771 AATGGGAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTAT 3830

Qy    650 PheLysLeuValAsnAlaGlnSer 657
Db   3831 TTTTCAATGGTCAGTGTCCAGGCT 3854
```

RESULT 3
AAZ88974
ID AAZ88974 standard; DNA; 3924 BP.
XX
AC AAZ88974;
XX
DT 30-MAY-2000 (first entry)
XX
DE Human MDR-3 DNA.
XX
KW Phosphatidylcholine; treatment; disease; multidrug resistance; MDR-3;
KW stomach; colon; ulcerative colitis; pouchitis; large intestine; human;
KW inflammation; Crohn's disease; diverse colitis; carcinoma; gene therapy;
KW ileum; anti-ulcer; anti-inflammatory; cytostatic; ss.
XX
OS Homo sapiens.
XX
PN WO200007577-A2.
XX
PD 17-FEB-2000.
XX
PF 06-AUG-1999; 99WO-DE02426.
XX
PR 06-AUG-1998; 98DE-1035526.
PR 15-DEC-1998; 98DE-1057750.
XX
PA (STRE/) STREMMEL W.

```
XX
PI   Stremmel W;
XX
DR   WPI; 2000-195439/17.
XX
PT   Using phosphatidylcholine having mucous membrane protective activity to
PT   prevent or treat large intestinal diseases, such as ulcerative colitis
PT   or pouchitis  -
XX
PS   Disclosure; Page 17-22; 22pp; German.
XX
CC   This invention describes a novel medicament containing a therapeutically
CC   effective amount of phosphatidylcholine to treat diseases, where
CC   phosphatidylcholine has an advantageous mucous membrane protective
CC   activity in the large intestine. The invention describes (1) a method
CC   to regulate presence or absence of MDR(multidrug resistance)-3
CC   transcription in a sample from patients with suspicion of stomach and
CC   colon diseases, where treatment with phosphatidylcholine is indicated;
CC   (2) a diagnostic test, to detect ulcerative colitis, pouchitis, large
CC   intestinal inflammation, Crohn's disease, diverse colitis, infectious
CC   enteritis/colitis, inflammation through X-ray treatment, antibiotics,
CC   chemotherapeutics, drugs or chemicals or large intestinal carcinoma,
CC   contains MDR3-specific primers to detect mutations in the MDR3 gene
CC   family and their interactions partners (in particular transcription
CC   factors); and (3) a method to manufacture a medicament for local gene
CC   therapy in the ileum with the ideal MDR3-analogous genes for patients
CC   with missing or decreased expression of MDR3-analogous proteins. The
CC   products of the invention have anti-ulcer, anti-inflammatory and
CC   cytostatic activity. The orally administered delayed-release form of
CC   phosphatidylcholine prevents premature resorption and provides for
CC   targeted release in the lower section of the small or large intestine.
CC   This sequence encodes the human MDR-3 protein described in the method of
CC   the invention.
XX
SQ   Sequence 3924 BP; 1145 A; 790 C; 977 G; 1012 T; 0 other;

Alignment Scores:
Pred. No.:              5.43e-187     Length:         3924
Score:                  1924.50       Matches:        372
Percent Similarity:     74.10%        Conservative:   123
Best Local Similarity:  55.69%        Mismatches:     162
Query Match:            57.91%        Indels:         11
DB:                     21            Gaps:           3

US-09-873-409-1 (1-659) x AAZ88974 (1-3924)

Qy        1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db     1851 GTAATTGTGGAGCAAGGAAGCCACAGCGAACTGATGAAGAAGGAAGGGGTGTACTTCAAA 1910

Qy       21 LeuVal---------------------------MetSerGlnAspIleLysLysAlaAsp 31
Db     1911 CTTGTCAACATGCAGACATCAGGAAGCCAGATCCAGTCAGAAGAATTTGAACTAAATGAT 1970

Qy       32 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 51
Db     1971 GAAAAGGCTGCCACTAGAATGGCCCCAAATGGCTGGAAATCTCGCCTATTTAGGCATTCT 2030

Qy       52 SerValLysSerIleLys---SerAspPheIleAspLysAla---GluGluSerThrGln 69
Db     2031 ACTCAGAAAAACCTTAAAAATTCACAAATGTGTCAGAAGAGCCTTGATGTGGAAACCGAT 2090

Qy       70 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 89
Db     2091 GGACTTGAAGCAAATGTGCCACCAGTGTCCTTTCTGAAGGTCCTGAAACTGAATAAAACA 2150

Qy       90 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 109
Db     2151 GAATGGCCCTACTTTGTCGTGGGAACAGTATGTGCCATTGCCAATGGGGGGCTTCAGCCG 2210

Qy      110 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 129
Db     2211 GCATTTTCAGTCATATTCTCAGAGATCATAGCGATTTTTGGACCAGGCGATGATGCAGTG 2270

Qy      130 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 149
Db     2271 AAGCAGCAGAAGTGCAACATATTCTCTTTGATTTTCTTATTTCTGGGAATTATTTCTTTT 2330

Qy      150 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 169
Db     2331 TTTACTTTCTTCCTTCAGGGTTTCACGTTTGGGAAAGCTGGCGAGATCCTCACCAGAAGA 2390

Qy      170 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLys 189
Db     2391 CTGCGGTCAATGGCTTTTAAAGCAATGCTAAGACAGGACATGAGCTGGTTTGATGACCAT 2450

Qy      190 GluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 209
Db     2451 AAAAACAGTACTGGTGCACTTTCTACAAGACTTGCCACAGATGCTGCCCAAGTCCAAGGA 2510

Qy      210 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 229
Db     2511 GCCACAGGAACCAGGTTGGCTTTAATTGCACAGAATATAGCTAACCTTGGAACTGGTATT 2570

Qy      230 IleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVal 249
Db     2571 ATCATATCATTTATCTACGGTTGGCAGTTAACCCTATTGCTATTAGCAGTTGTTCCAATT 2630

Qy      250 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 269
Db     2631 ATTGCTGTGTCAGGAATTGTTGAAATGAAATTGTTGGCTGGAAATGCCAAAAGAGATAAA 2690

Qy      270 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIle 289
Db     2691 AAAGAACTGGAAGCTGCTGGAAAGATTGCAACAGAGGCAATAGAAAATATTAGGACAGTT 2750

Qy      290 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGln 309
Db     2751 GTGTCTTTGACCCAGGAAAGAAAATTTGAATCAATGTATGTTGAAAAATTGTATGGACCT 2810

Qy      310 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla 329
Db     2811 TACAGGAATTCTGTGCAGAAGGCACACATCTATGGAATTACTTTTAGTATCTCACAAGCA 2870

Qy      330 PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGly 349
Db     2871 TTTATGTATTTTTCCTATGCCGGTTGTTTTCGATTTGGTGCATATCTCATTGTGAATGGA 2930

Qy      350 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAla 369
Db     2931 CATATGCGCTTCAGAGATGTTATTCTGGTGTTTTCTGCAATTGTATTTGGTGCAGTGGCT 2990

Qy      370 IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHis 389
Db     2991 CTAGGACATGCCAGTTCATTTGCTCCAGACTATGCTAAAGCTAAGCTGTCTGCAGCCCAC 3050

Qy      390 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLys 409
Db     3051 TTATTCATGCTGTTTGAAAGACAACCTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAG 3110

Qy      410 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 429
Db     3111 CCTGATAAATTTGAAGGAAATATAACATTTAATGAAGTCGTGTTCAACTATCCCACCCGA 3170

Qy      430 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 449
Db     3171 GCAAACGTGCCAGTGCTTCAGGGGCTGAGCCTGGAGGTGAAGAAAGGCCAGACACTAGCC 3230

Qy      450 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr 469
Db     3231 CTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACGGTGGTCCAGCTCCTGGAGCGGTTCTAC 3290

Qy      470 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 489
Db     3291 GACCCCTTGGCGGGGACAGTGCTTCTCGATGGTCAAGAAGCAAAGAAACTCAATGTCCAG 3350
```

```
Qy     490 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 509
Db    3351 TGGCTCAGAGCTCAACTCGGAATCGTGTCTCAGGAGCCTATCCTATTTGACTGCAGCATT 3410

Qy     510 AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu 529
Db    3411 GCCGAGAATATTGCCTATGGAGACAACAGCCGGGTTGTATCACAGGATGAAATTGTGAGT 3470

Qy     530 AlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThr 549
Db    3471 GCAGCCAAAGCTGCCAACATACATCCTTTCATCGAGACGTTACCCCACAAATATGAAACA 3530

Qy     550 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAla 569
Db    3531 AGAGTGGGAGATAAGGGGACTCAGCTCTCAGGAGGTCAAAAACAGAGGATTGCTATTGCC 3590

Qy     570 ArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsp 589
Db    3591 CGAGCCCTCATCAGACAACCTCAAATCCTCCTGTTGGATGAAGCTACATCAGCTCTGGAT 3650

Qy     590 AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys 609
Db    3651 ACTGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGC 3710

Qy     610 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 629
Db    3711 ATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAG 3770

Qy     630 AsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyr 649
Db    3771 AATGGGAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTAT 3830

Qy     650 PheLysLeuValAsnAlaGlnSer 657
Db    3831 TTTCAATGGTCAGTGTCCAGGCT 3854
```

RESULT 4
ABN95801

```
ID   ABN95801 standard; DNA; 3924 BP.
XX
AC   ABN95801;
XX
DT   13-AUG-2002  (first entry)
XX
DE   Gene #2299 used to diagnose liver cancer.
XX
KW   Gene; liver cancer; ds; hepatocellular carcinoma; hepatotropic;
KW   metastatic liver tumour; cytostatic; expression profile; disease state;
KW   disease progression; drug toxicity; drug efficacy; drug metabolism.
XX
OS   Homo sapiens.
XX
PN   WO200229103-A2.
XX
PD   11-APR-2002.
XX
PF   02-OCT-2001; 2001WO-US30589.
XX
PR   02-OCT-2000; 2000US-237054P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Horne D,  Alvares C,  Peres-Da-Silva S,  Vockley JG;
XX
DR   WPI; 2002-426119/45.
XX
PT   Diagnosing and detecting the progression of liver cancer,
PT   hepatocellular carcinoma or metastatic liver tumor in a patient,
PT   involves detecting the level of expression of two or more genes in a
PT   liver tissue sample  -
XX
PS   Claim 1; SEQ ID NO 2299; 298pp; English.
XX
CC   The invention relates to a novel method for diagnosing and detecting the
CC   progression of liver cancer, hepatocellular carcinoma or metastatic liver
CC   tumour in a patient, and differentiating metastatic liver cancer from
CC   hepatocellular carcinoma in a patient, involving detecting the level of
CC   expression of two or more genes represented in ABN93503-ABN97455 in a
CC   tissue sample. The method of the invention has hepatotropic, and
CC   cytostatic activity. The method is useful for diagnosing and detecting
CC   the progression of liver cancer, hepatocellular carcinoma and metastatic
CC   liver carcinoma in a patient. The method is useful for identifying
CC   expression profiles which serve as useful diagnostic markers as well as
CC   markers that can be used to monitor disease states, disease progression,
CC   drug toxicity, drug efficacy and drug metabolism.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic format directly from WIPO
CC   at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 3924 BP; 1145 A; 790 C; 977 G; 1012 T; 0 other;
```

```
Alignment Scores:
Pred. No.:              5.43e-187      Length:        3924
Score:                  1924.50        Matches:       372
Percent Similarity:     74.10%         Conservative:  123
Best Local Similarity:  55.69%         Mismatches:    162
Query Match:            57.91%         Indels:        11
DB:                     24             Gaps:          3
```

US-09-873-409-1 (1-659) x ABN95801 (1-3924)

```
Qy       1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db    1851 GTAATTGTGGAGCAAGGAAGCCACAGCGAACTGATGAAGAAGGAAGGGGTGTACTTCAAA 1910

Qy      21 LeuVal---------------------------MetSerGlnAspIleLysLysAlaAsp 31
Db    1911 CTTGTCAACATGCAGACATCAGGAAGCCAGATCCAGTCAGAAGAATTTGAACTAAATGAT 1970

Qy      32 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 51
Db    1971 GAAAAGGCTGCCACTAGAATGGCCCCAAATGGCTGGAAATCTCGCCTATTTAGGCATTCT 2030

Qy      52 SerValLysSerIleLys---SerAspPheIleAspLysAla---GluGluSerThrGln 69
Db    2031 ACTCAGAAAAACCTTAAAAATTCACAAATGTGTCAGAAGAGCCTTGATGTGGAAACCGAT 2090

Qy      70 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 89
Db    2091 GGACTTGAAGCAAATGTGCCACCAGTGTCCTTTCTGAAGGTCCTGAAACTGAATAAAACA 2150

Qy      90 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 109
Db    2151 GAATGGCCCTACTTTGTCGTGGGAACAGTATGTGCCATTGCCAATGGGGGGCTTCAGCCG 2210

Qy     110 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 129
Db    2211 GCATTTTCAGTCATATTCTCAGAGATCATAGCGATTTTTGGACCAGGCGATGATGCAGTG 2270

Qy     130 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 149
Db    2271 AAGCAGCAGAAGTGCAACATATTCTCTTTGATTTTCTTATTTCTGGGAATTATTTCTTTT 2330

Qy     150 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 169
Db    2331 TTTACTTTCTTCCTTCAGGGTTTCACGTTTGGGAAAGCTGGCGAGATCCTCACCAGAAGA 2390

Qy     170 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLys 189
Db    2391 CTGCGGTCAATGGCTTTTAAAGCAATGCTAAGACAGGACATGAGCTGGTTTGATGACCAT 2450

Qy     190 GluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 209
Db    2451 AAAAACAGTACTGGTGCACTTTCTACAAGACTTGCCACAGATGCTGCCCAAGTCCAAGGA 2510

Qy     210 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 229
```

```
                |||||||||:::|||:::    ::::::  ||||||       |||:::|||       :::
Db       2511 GCCACAGGAACCAGGTTGGCTTTAATTGCACAGAATATAGCTAACCTTGGAACTGGTATT 2570

Qy        230 IleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVal 249
              |||||||||||||||||||||||||::::::|||   |||:::|||::::::   |||:::
Db       2571 ATCATATCATTTATCTACGGTTGGCAGTTAACCCTATTGCTATTAGCAGTTGTTCCAATT 2630

Qy        250 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 269
              :::||||||:::|||::::::|||          :::   |||   |||   :::||||||
Db       2631 ATTGCTGTGTCAGGAATTGTTGAAATGAAATTGTTGGCTGGAAATGCCAAAAGAGATAAA 2690

Qy        270 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIle 289
              :::||||||:::   |||||||||||||||||||||||||:::||||||||||||||||:::
Db       2691 AAAGAACTGGAAGCTGCTGGAAAGATTGCAACAGAGGCAATAGAAAATATTAGGACAGTT 2750

Qy        290 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGln 309
              ||||||||||||:::|||:::   ||||||   ||||||   |||   |||
Db       2751 GTGTCTTTGACCCAGGAAAGAAAATTTGAATCAATGTATGTTGAAAAATTGTATGGACCT 2810

Qy        310 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla 329
              :::||||||:::   :::||||||   |||   |||      ::::::    |||   |||
Db       2811 TACAGGAATTCTGTGCAGAAGGCACACATCTATGGAATTACTTTTAGTATCTCACAAGCA 2870

Qy        330 PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGly 349
              |||:::||||||:::||||||      |||||||||||||||||||||   |||
Db       2871 TTTATGTATTTTTCCTATGCCGGTTGTTTTCGATTTGGTGCATATCTCATTGTGAATGGA 2930

Qy        350 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAla 369
                 |||            :::   :::||||||:::||||||   :::||||||:::|||
Db       2931 CATATGCGCTTCAGAGATGTTATTCTGGTGTTTTCTGCAATTGTATTTGGTGCAGTGGCT 2990

Qy        370 IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHis 389
              :::|||                  ||||||:::|||:::|||||||||      |||||||||
Db       2991 CTAGGACATGCCAGTTCATTTGCTCCAGACTATGCTAAAGCTAAGCTGTCTGCAGCCCAC 3050

Qy        390 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLys 409
              ||||||   |||   |||::::::|||   |||||||||   |||:::||||||   |||
Db       3051 TTATTCATGCTGTTTGAAAGACAACCTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAG 3110

Qy        410 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 429
              ||||||      ||||||||:::   |||   ||||||   |||   ||||||   |||
Db       3111 CCTGATAAATTTGAAGGAAATATAACATTTAATGAAGTCGTGTTCAACTATCCCACCCGA 3170

Qy        430 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 449
                 :::|||   :::|||:::||||||||||||   :::::::::|||:::|||:::|||
Db       3171 GCAAACGTGCCAGTGCTTCAGGGGCTGAGCCTGGAGGTGAAGAAAGGCCAGACACTAGCC 3230

Qy        450 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr 469
                 |||||||||||||||||||||||||||||   ||||||||||||:::|||   |||
Db       3231 CTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACGGTGGTCCAGCTCCTGGAGCGGTTCTAC 3290

Qy        470 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 489
              ||||||:::   |||   ||||||   ||||||   :::||||||:::||||||||||||
Db       3291 GACCCCTTGGCGGGGACAGTGCTTCTCGATGGTCAAGAAGCAAAGAAACTCAATGTCCAG 3350

Qy        490 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 509
              |||||||||:::|||:::         ||||||   ||||||||||||:::||||||||
Db       3351 TGGCTCAGAGCTCAACTCGGAATCGTGTCTCAGGAGCCTATCCTATTTGACTGCAGCATT 3410

Qy        510 AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu 529
              ||||||||||||||||||||||||||||||||||||         |||||||||
Db       3411 GCCGAGAATATTGCCTATGGAGACAACAGCCGGGTTGTATCACAGGATGAAATTGTGAGT 3470

Qy        530 AlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThr 549
              ||||||   |||||||||||||||   |||||||||   ||||||   ||||||   |||
Db       3471 GCAGCCAAAGCTGCCAACATACATCCTTTCATCGAGACGTTACCCCACAAATATGAAACA 3530

Qy        550 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAla 569
              :::||||||   ||||||   |||||||||||||||||||||||||||:::|||||||||
Db       3531 AGAGTGGGAGATAAGGGGACTCAGCTCTCAGGAGGTCAAAAACAGAGGATTGCTATTGCC 3590

Qy        570 ArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsp 589
              ||||||||:::::::::|||:::||||||||||||||||||||||||||||||||||||
Db       3591 CGAGCCCTCATCAGACAACCTCAAATCCTCCTGTTGGATGAAGCTACATCAGCTCTGGAT 3650

Qy        590 AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys 609
                 :::||||||||||||||||||   ||||||||||||||||||   ||||||||||||
Db       3651 ACTGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGC 3710

Qy        610 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 629
                 :::|||:::      ||||||||||||   |||||||||||||||||||||||||||
Db       3711 ATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAG 3770

Qy        630 AsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyr 649
              ||||||::::::||||||   |||||||||||:::||||||         :::   ||||||
Db       3771 AATGGGAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTAT 3830

Qy        650 PheLysLeuValAsnAlaGlnSer 657
              |||   :::|||:::   |||:::
Db       3831 TTTCAATGGTCAGTGTCCAGGCT 3854
```

```
RESULT 5
ABK63653
ID   ABK63653 standard; cDNA; 3912 BP.
XX
AC   ABK63653;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Rat sequence differentially expressed in response to a hepatotoxin #1560.
XX
KW   Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
KW   differential expression; centrilobular necrosis; steatosis.
XX
OS   Rattus norvegicus.
XX
PN   WO200210453-A2.
XX
PD   07-FEB-2002.
XX
PF   30-JUL-2001; 2001WO-US23872.
XX
PR   31-JUL-2000; 2000US-222040P.
PR   02-NOV-2000; 2000US-244880P.
PR   11-MAY-2001; 2001US-290029P.
PR   15-MAY-2001; 2001US-290645P.
PR   22-MAY-2001; 2001US-292336P.
PR   06-JUN-2001; 2001US-295798P.
PR   13-JUN-2001; 2001US-297457P.
PR   19-JUN-2001; 2001US-298884P.
PR   09-JUL-2001; 2001US-303459P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Mendrick D,  Porter MW,  Johnson KR,  Castle AL,  Elashoff MR;
XX
DR   WPI; 2002-241625/29.
XX
PT   Predicting toxic effects of compounds or the progression of these toxic
PT   effects by determining the changes in gene expression in tissues or
PT   cells exposed to the toxin and comparing these to gene expression in
PT   unexposed tissues or cells -
XX
PS   Claim 1; Seq ID No 1560; 239pp; English.
XX
CC   The invention relates to methods for predicting toxic effects of
CC   compounds or the progression of these toxic effects by determining the
CC   global changes in gene expression in tissues or cells exposed to the
CC   toxin and comparing these to gene expression in unexposed tissues or
CC   cells. Also included are methods of predicting at least one toxic
CC   effect of a compound or progression of a toxic effect, preferably the
CC   hepatotoxicity of a compound, comprising detecting the level of
CC   expression in a tissue or cell sample exposed to the compound of two or
CC   more genes listed in the specification, where differential expression of
CC   the genes is indicative of at least one toxic effect or progression.
CC   The method can also be used to identify an agent which modulates the
```

```
CC   toxic response and predict cellular pathways that a compound modulates
CC   in a cell. The methods utilise a set of at least two probes (on a solid
CC   support in kit form), where each of the probes comprises a sequence that
CC   specifically hybridises to a gene listed in the specification, a computer
CC   system comprising a database containing information identifying the
CC   expression level in a tissue or cell sample exposed to a hepatotoxin of a
CC   set of genes comprising at least two genes listed in the specification,
CC   and a user interface to view the information used to present information
CC   identifying the expression level in a tissue or cell of at least one gene
CC   listed in the specification. The method is useful for elucidating global
CC   changes in gene expression and for identifying toxicity markers in
CC   tissues or cell exposed to a known toxin. The genes may be used as
CC   toxicity markers in drug screening and toxicity assays. The genes and
CC   gene expression information may be used as diagnostic markers for the
CC   prediction or identification of the physiological state of tissue or cell
CC   sample that has been exposed to a compound or agent. Hepatotoxicity
CC   is characterised by centrilobular necrosis and steatosis. The present
CC   sequence is an expressed sequence tag (EST) or cDNA derived from a gene
CC   which is differentially expressed in response to a hepatotoxic agent.
XX
SQ   Sequence 3912 BP; 1077 A; 904 C; 1017 G; 914 T; 0 other;

Alignment Scores:
Pred. No.:              1.12e-182      Length:         3912
Score:                  1882.50        Matches:        363
Percent Similarity:     73.43%         Conservative:   129
Best Local Similarity:  54.18%         Mismatches:     165
Query Match:            56.65%         Indels:         13
DB:                     24             Gaps:           2

US-09-873-409-1 (1-659) x ABK63653 (1-3912)

Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
          ::::::    |||:::|||:::|||:::|||||:::    |||    |||:::|||:::
Db   1842 GTCATCGTGGAGCAAGGAAGCCACAGTGAGCTGATAAAGAAGGAAGGGATCTACTTCAGA 1901

Qy     21 LeuValMetSerGln--------------------------------AspIleLysLys 29
          ||||||    |||                                   ::::::::
Db   1902 CTTGTTAACATGCAGACATCAGGAAGCCAGATCCTGTCAGAAGAATTTGAAGTTGAGCTA 1961

Qy     30 AlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPro 49
          :::||||||:::                                 |||
Db   1962 AGTGATGAAAAGGCTGCTGGAGGTGTGGCCCCAAATGGCTGGAAAGCACGCATATTTAGG 2021

Qy     50 LeuHisSerValLysSerIleLysSerAsp------PheIleAspLysAlaGluGluSer 67
             :::   ||||||:::||||||              ::::::   :::
Db   2022 AATTCTACGAAGAAAAGTCTGAAAAGTTCACGGGCGCATCAAAATAGGCTGGATGTGGAA 2081

Qy     68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
          |||         :::   ::::::|||   ||||||   ||||||:::|||:::||||||
Db   2082 ACCAATGAACTTGATGCAAACGTGCCACCAGTGTCTTTTCTGAAGGTCTTAAGACTGAAT 2141

Qy     88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
          |||   |||||||||:::   |||:::|||||||||   ::::::   ||||||   :::
Db   2142 AAAACAGAGTGGCCCTACTTTGTGGTGGGGACACTCTGTGCCATTGCCAACGGGGCCCTC 2201

Qy    108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 127
             |||   ||||||||||||   :::::::::|||   :::||||||      |||
Db   2202 CAGCCGGCATTCTCCATCATCCTGTCAGAGATGATAGCTATCTTTGGCCCTGGGGATGAC 2261

Qy    128 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 147
          |||      :::         ::::::|||::::::|||:::   |||||||||
Db   2262 ACAGTAAAGCAACAGAAGTGTAACATGTTCTCGCTGGTCTTCTTGGGCCTAGGAGTCCAC 2321

Qy    148 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 167
             |||   ::::::|||:::||||||      :::|||:::||||||||||||||||||
Db   2322 TCCTTCTTTACTTTCTTCCTTCAGGGTTTCACATTCGGGAAAGCTGGCGAGATCCTCACC 2381

Qy    168 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 187
             |||||||||   :::|||||||||||||||||||   ||||||::::::|||||||||
Db   2382 ACAAGGCTCCGGTCCATGGCCTTCAAAGCAATGCTAAGACAGGACATGAGCTGGTTTGAC 2441

Qy    188 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 207
          :::   :::|||||||||||   |||:::|||   ||||||   |||   ||||||:::
Db   2442 GATCATAAAAACAGTACTGGTGCCCTCTCTACAAGACTCGCCACAGACGCTGCGCAGGTC 2501

Qy    208 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 227
          ||||||||||||||||||:::|||:::      ::::::   ||||||      |||:::|||
Db   2502 CAAGGAGCCACAGGAACCAGGTTGGCTTTAATTGCACAGAACACAGCCAACCTTGGAACG 2561

Qy    228 SerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAla 247
             :::||||||||||||||||||||||||||||::::::|||   |||:::||||||:::
Db   2562 GGTATTATTATATCATTTATTTACGGTTGGCAACTGACACTTCTGCTCTTATCAGTTGTT 2621

Qy    248 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 267
          |||   :::||||||   |||:::::::|||         :::   |||   |||   :::
Db   2622 CCATTCATTGCTGTAGCGGGAATTGTTGAAATGAAAATGTTGGCTGGCAACGCCAAGAGA 2681

Qy    268 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 287
          ||||||:::|||::::::   |||||||||||||||||||||||||:::||||||||||||
Db   2682 GATAAAAAGGAGATGGAAGCTGCTGGAAAGATTGCAACAGAGGCAATAGAAAATATTCGG 2741

Qy    288 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 307
          |||:::||||||||||||:::|||:::      ||||||   ||||||   |||   |||
Db   2742 ACTGTTGTATCCTTGACCCAAGAGAGAAAATTTGAGTCAATGTATGTTGAAAAATTACAC 2801

Qy    308 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 327
                :::|||||||||:::   :::||||||   |||         ::::::   |||
Db   2802 GGACCTTACAGGAATTCAGTGCGGAAGGCTCACATCTACGGCATCACTTTTAGCATCTCA 2861

Qy    328 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 347
             ||||||:::||||||:::||||||      ||||||||||||:::|||||||||
Db   2862 CAAGCATTCATGTACTTTTCTTATGCTGGCTGCTTTCGATTTGGTTCTTACCTCATTGTG 2921

Qy    348 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 367
             |||   |||      :::   :::   :::|||||||:::||||||      ||||||
Db   2922 AATGGACACATGCGCTTCAAGGATGTCATCCTGGTGTTCTCAGCAATCGTGCTTGGTGCA 2981

Qy    368 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 387
          :::|||:::|||               ||||||:::|||:::|||||||||      |||
Db   2982 GTGGCTCTAGGACATGCCAGCTCATTTGCTCCAGACTATGCAAAAGCCAAGCTGTCTGCA 3041

Qy    388 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 407
          |||:::|||||||||:::|||   |||||||||:::::|||      |||   |||:::||||||
Db   3042 GCATACTTATTCAGTCTGTTTGAAAGACAACCTCTGATTGACAGCTACAGCAGAGAAGGA 3101

Qy    408 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 427
                ||||||      ||||||::::::   |||   ||||||   |||   ||||||
Db   3102 ATGTGGCCGGATAAGTTTGAAGGAAGCGTGACATTCAATGAAGTTGTGTTCAATTATCCC 3161

Qy    428 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 447
             |||   :::|||   :::|||:::|||||||||||||   :::::::::|||:::|||
Db   3162 ACCCGGGCCAATGTGCCAGTGCTTCAGGGGCTGAGCCTCGAGGTGAAGAAGGGGCAGACC 3221

Qy    448 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 467
          :::|||   ||||||||||||||||||||||||||||||   ||||||||||||:::|||
Db   3222 CTGGCCCTGGTGGGCAGTAGTGGCTGCGGGAAGAGCACCGTGGTCCAGCTGCTCGAGCGC 3281

Qy    468 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 487
          |||||||||:::   |||   ||||||   ||||||   :::||||||:::||||||
Db   3282 TTCTACGACCCCATGGCCGGAACAGTGCTCCTCGATGGTCAGGAAGCAAAGAAACTCAAT 3341

Qy    488 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 507
          |||||||||||||||:::|||:::   ||||||   |||||||||:::||||||:::|||
Db   3342 GTCCAGTGGCTCCGAGCTCAACTTGGCATTGTGTCCCAGGAGCCCATCCTGTTTGACTGC 3401

Qy    508 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 527
          |||||||||:::|||||||||||||||||||||||||||||||||   |||||||||
Db   3402 AGCATCGCCAAGAACATCGCCTACGGAGACAACAGCCGTGTCGTGTCTCAGGATGAGATT 3461

Qy    528 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 547
                ||||||      ||||||||||||   |||||||||   ||||||:::||||||
Db   3462 GTGAGGGCGGCCAAGGAGGCCAACATCCACCCCTTCATTGAGACACTGCCCCAAAAGTAT 3521

Qy    548 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 567
          |||:::||||||   ||||||   ||||||||||||||||||||||||:::|||
```

```
Db   3522 GAAACAAGAGTAGGAGACAAGGGGACACAGCTCTCTGGAGGCCAGAAACAGAGGATTGCT 3581

Qy    568 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 587

Db   3582 ATCGCCCGAGCCCTCATCAGACAGCCTCGGGTCCTACTGCTGGATGAAGCCACGTCGGCT 3641

Qy    588 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 607

Db   3642 TTGGACACTGAGAGTGAAAAGGTCGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGC 3701

Qy    608 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 627

Db   3702 ACCTGCATTGTGATCGCGCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTG 3761

Qy    628 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 647

Db   3762 ATCGACAACGGCAAGGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCCCAGAAAGGC 3821

Qy    648 IleTyrPheLysLeuValAsnAlaGlnSer 657

Db   3822 ATCTATTTCTCCATGGTCAACATTCAAGCT 3851
```

RESULT 6
AAZ52047
ID   AAZ52047 standard; cDNA; 4369 BP.
XX
AC   AAZ52047;
XX
DT   18-JUL-2000 (first entry)
XX
DE   Rat multidrug resistance protein 1a cDNA.
XX
KW   Multidrug resistance protein 1a; mdr1a; multi-specific drug transporter;
KW   drug formulation; formulation excipient; compound design; inflammation;
KW   pharmacokinetic; oral absorption; cancer; cardiovascular disease;
KW   central nervous system disorder; auto-immune disease; kidney disease; ss.
XX
OS   Rattus rattus.
XX
FH   Key   Location/Qualifiers
FT   CDS   352..4170
FT   /*tag= a
FT   /product= "Rat multidrug resistance protein"
XX
PN   WO200015650-A1.
XX
PD   23-MAR-2000.
XX
PF   10-SEP-1999; 99WO-US20770.
XX
PR   17-SEP-1998; 98US-0156800.
PR   09-DEC-1998; 98US-0208809.
XX
PA   (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI   Brun KA, Chenery RJ, Ellens H, Feild JA, Yue L;
XX
DR   WPI; 2000-271372/23.
DR   P-PSDB; AAY70596.
XX
PT   Isolated rat mdr1a polynucleotides and polypeptides, useful in assays
PT   to provide information on drug formulation, selection of formulation
PT   excipients and compound design -
XX
PS   Claim 2; Page 23-24; 33pp; English.
XX
CC   The present cDNA sequence encodes rat multidrug resistance
CC   protein 1a (mdr1a). This protein is a member of the multi-specific drug
CC   transporters family. Mdr1a is used in assays to provide information on
CC   drug formulation, selection of formulation excipients and compound
CC   design. They are used in cell based, membrane based, binding or other
CC   assays to provide information that may enhance drug formulation. This
CC   invention further relates to the generation of in vivo and in vitro
CC   comparison data to predict oral absorption and pharmacokinetics. This
CC   enables the selection of drugs with optimal pharmacokinetics, i.e. good
CC   oral bioavailability, brain penetration, plasma half life, and minimum
CC   drug interaction. Transgenic and knock-out animals created using rat
CC   mdr1a provides an insight into treating and preventing human diseases
CC   including cancer, inflammation, cardiovascular disease, central nervous
CC   system disorders, auto-immune and kidney disease.
XX
SQ   Sequence 4369 BP; 1241 A; 973 C; 1096 G; 1059 T; 0 other;

Alignment Scores:			
Pred. No.:	1.49e-182	Length:	4369
Score:	1882.00	Matches:	362
Percent Similarity:	74.37%	Conservative:	137
Best Local Similarity:	53.95%	Mismatches:	158
Query Match:	56.64%	Indels:	14
DB:	21	Gaps:	5

US-09-873-409-1 (1-659) x AAZ52047 (1-4369)

```
Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20

Db   2140 GTCATTGTGGAGCAAGGAAATCATGATGAGCTCATGAGAGAGAAAGGAATTTACTTCAAA 2199

Qy     21 LeuValMetSerGln------------AspIleLysLysAlaAspGluGlnMetGluSer 36

Db   2200 CTTGTCATGACTCAGACAGCAGGAAATGAAATTGAATTAGGAAATGAAGCTTGTGAATCT 2259

Qy     37 ------------------MetThrTyrSerThrGluArgLysThrAsnSerLeu------ 48

Db   2260 AAAGACGGAATTGATAATGTGGACATGTCTTCAAAAGATTCGGGATCCAGTCTAATAAGA 2319

Qy     49 ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr 68

Db   2320 AGAAGATCAACTCGCAAAAGCATCCGTGGGCCACATGATCAAGACGGGGAACTTAGCACC 2379

Qy     69 ---GlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87

Db   2380 AAAGAGGCTCTGGATGACGACGTACCTCCAGCTTCCTTTTGGCGGATCCTGAAGTTGAAT 2439.

Qy     88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107

Db   2440 TCAACTGAATGGCCTTATTTTGTGGTTGGTGTATTTTGTGCCATAATAAATGGAGGCTTG 2499

Qy    108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 127

Db   2500 CAACCAGCATTCTCCATAATATTTTCAAAGGTTGTAGGGGTTTTTACAAAAAATGACACC 2559

Qy    128 ThrThrLeu---LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146

Db   2560 CCTGAAATCCAGCGGCAGAACAGCAACTTGTTTTCTTTATTGTTCCTGATCCTTGGGATC 2619

Qy    147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166

Db   2620 ATCTCTTTCATTACGTTTTTCCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTC 2679

Qy    167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186

Db   2680 ACCAAGCGACTCCGATACATGGTCTTCAAATCCATGCTGAGACAGGACATAAGCTGGTTT 2739

Qy    187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206

Db   2740 GATGACCCTAAAAACACCACAGGAGCGCTGACCACCAGGCTTGCCAATGACGCTGCTCAA 2799

Qy    207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226

Db   2800 GTGAAAGGGGCTACAGGGTCTAGGCTTGCTGTTATTACCCAGAACATAGCAAATCTTGGG 2859

Qy    227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246

Db   2860 ACAGGCATCATCATATCCCTGATCTACGGCTGGCAATTGACACTTTTACTCCTAGCAATT 2919

Qy    247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
```

```
            |||:::::|||:::   |||:::::|||          ::::::|||   |||
Db    2920 GTTCCCATCATTGCTATAGCAGGAGTGGTTGAAATGAAAATGTTGTCTGGACAAGCGCTG 2979

Qy     267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
           |||||||||:::||||||:::   :::||||||||||||||||||||:::||||||
Db    2980 AAAGATAAGAAGGAACTAGAAGGTTCTGGGAAGATCGCTACAGAAGCAATTGAAAACTTT 3039

Qy     287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
           ||||||:::||||||||||||||||||:::   ||||||   ||||||   :::   |||
Db    3040 CGCACTGTCGTCTCTTTGACTCGGGAGCAGAAGTTTGAAACTATGTATGCCCAGAGCTTG 3099

Qy     307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
           |||         :::||||||   ||||||||   :::   |||      ::::::|||
Db    3100 CAGATACCATACAGAAATGCTTTGAAGAAAGCGCACGTCTTTGGGATCACTTTCTCCTTC 3159

Qy     327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
           :::   |||   :::||||||:::||||||||   |||||||||   |||||||||:::
Db    3160 ACCCAGGCCATGATGTATTTCTCCTATGCTGCTTGTTTCCGGTTTGATGCCTACTTGGTG 3219

Qy     347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
                          ||||||   |||   :::   :::||||||:::||||||   :::|||
Db    3220 GCACGAGAACTCATGACATTTGAAAATGTTCTGTTAGTATTCTCAGCTATTGTCTTTGGT 3279

Qy     367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
           |||||||||:::|||:::                     ||||||:::|||:::||||||||
Db    3280 GCCATGGCAGTGGGGCAGGTCAGTTCATTCGCTCCTGACTACGCGAAAGCCAAAGTCTCG 3339

Qy     387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
           |||:::|||:::      ::::::||||||   |||   |||||||||   |||   |||
Db    3340 GCATCCCACATCATCAGGATCATTGAGAAAATCCCTGAGATTGACAGCTACAGCACGGAG 3399

Qy     407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
           |||   ||||||:::      |||||||||::::::|||      |||   |||   |||
Db    3400 GGCTTGAAGCCTAATATGTTGGAAGGAAATGTGAAATTTAATGGAGTCATGTTCAACTAT 3459

Qy     427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
           |||   ||||||::::::   :::|||:::|||||||||   :::::::::|||:::
Db    3460 CCCACCCGACCCAACATCCCAGTGCTTCAGGGGCTGAGCCTAGAGGTGAAGAAAGGGCAG 3519

Qy     447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
           |||:::|||   ||||||||||||||||||||||||||||||   ||||||||||||:::
Db    3520 ACGCTGGCCCTCGTGGGCAGCAGTGGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTTGAG 3579

Qy     467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
           |||   ||||||||||:::   |||   |||   ||||||   :::   |||:::|||
Db    3580 CGCTTCTATGACCCCATGGCCGGAACAGTGTTTCTAGATGGCAAAGAAATAAAGCAACTC 3639

Qy     487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
           ||||||||||||||||||:::   :::   ||||||   |||||||||:::||||||:::
Db    3640 AATGTCCAGTGGCTCCGCGCCCACCTGGGCATTGTGTCCCAGGAGCCCATCCTGTTTGAC 3699

Qy     507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
           ||||||||||||||||||||||||||||||||||||||||||||||||         |||
Db    3700 TGCAGCATCGCCGAGAACATTGCCTACGGAGACAACAGCCGTGTCGTGTCTCATAAGGAG 3759

Qy     527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
           |||   :::||||||      ||||||||||||   ||||||:::   ||||||||||||
Db    3760 ATCGTGAAGGCAGCCAAGGAGGCCAACATCCACCAGTTCATCGACTCACTGCCTGAGAAA 3819

Qy     547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
           |||||||||:::||||||   ||||||   |||||||||||||||||||||||||||:::
Db    3820 TACAACACCAGAGTGGGAGACAAAGGGACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATC 3879

Qy     567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
           ||||||||||||||||||:::::::::|||   |||||||||||||||||||||||||||
Db    3880 GCCATCGCGCGCGCCCTCGTCAGACAGCCTCACATCTTACTTCTGGATGAAGCGACATCA 3939

Qy     587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
           ||||||||   :::||||||||||||||||||   ||||||||||||||||||   |||
Db    3940 GCTCTGGATACGGAGAGTGAAAAGGTCGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGC 3999

Qy     607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
           |||||||||:::|||:::   ||||||||||||   ||||||||||||||||||||||||
Db    4000 CGCACCTGCATTGTGATCGCGCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTG 4059

Qy     627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
           |||:::   ||||||::::::||||||   ||||||||||||:::||||||      :::
Db    4060 GTGATTCAGAACGGCCAGGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCCCAGAAA 4119

Qy     647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
           |||||||||   :::|||:::   |||:::
Db    4120 GGCATCTATTTCTCGATGGTCAGTGTGCAGGCT 4152
```

```
RESULT 7
AAZ52048
ID   AAZ52048 standard; cDNA; 4425 BP.
XX
AC   AAZ52048;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Rat multidrug resistance protein 1a cDNA derived from EST sequences.
XX
KW   Multidrug resistance protein 1a; mdr1a; multi-specific drug transporter;
KW   drug formulation; formulation excipient; compound design; inflammation;
KW   pharmacokinetic; oral absorption; cancer; cardiovascular disease;
KW   central nervous system disorder; auto-immune disease; kidney disease;
KW   EST; expressed sequence tag; ss.
XX
OS   Rattus rattus.
XX
FH   Key             Location/Qualifiers
FT   CDS             352..4170
FT                   /*tag=  a
FT                   /product= "Rat multidrug resistance protein"
XX
PN   WO200015650-A1.
XX
PD   23-MAR-2000.
XX
PF   10-SEP-1999;   99WO-US20770.
XX
PR   17-SEP-1998;   98US-0156800.
PR   09-DEC-1998;   98US-0208809.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI   Brun KA,  Chenery RJ,  Ellens H,  Feild JA,  Yue L;
XX
DR   WPI; 2000-271372/23.
DR   P-PSDB; AAY70597.
XX
PT   Isolated rat mdr1a polynucleotides and polypeptides, useful in assays
PT   to provide information on drug formulation, selection of formulation
PT   excipients and compound design -
XX
PS   Claim 11; Page 27-29; 33pp; English.
XX
CC   The present cDNA sequence encodes rat multidrug resistance protein 1a
CC   (mdr1a). This cDNA is derived from EST (expressed sequence tag)
CC   sequences. Mdr1a is used in assays to provide information on
CC   drug formulation, selection of formulation excipients and compound
CC   design. They are used in cell based, membrane based, binding or other
CC   assays to provide information that may enhance drug formulation. This
CC   invention further relates to the generation of in vivo and in vitro
CC   comparision data to predict oral absorption and pharmacokinetics. This
CC   enables the selection of drugs with optimal pharmacokinetics, i.e. good
CC   oral bioavailability, brain penetration, plasma half life, and minimum
CC   drug interaction. Transgenic and knock-out animals created using rat
CC   mdr1a provides an insight into treating and preventing human diseases
CC   including cancer, inflammation, cardiovascular disease, central nervous
CC   system disorders, auto-immune and kidney disease.
XX
SQ   Sequence 4425 BP; 1250 A; 992 C; 1113 G; 1067 T; 2 U; 1 other;

Alignment Scores:
```

```
Pred. No.:               1.52e-182     Length:         4425
Score:                   1882.00       Matches:        362
Percent Similarity:      74.37%        Conservative:   137
Best Local Similarity:   53.95%        Mismatches:     158
Query Match:             56.64%        Indels:         14
DB:                      21            Gaps:           5

US-09-873-409-1 (1-659) x AAZ52048 (1-4425)

Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db   2140 GTCATTGTGGAGCAAGGAAATCATGATGAGCTCATGAGAGAGAAAGGAATTTACTTCAAA 2199

Qy     21 LeuValMetSerGln-----------AspIleLysLysAlaAspGluGlnMetGluSer 36
Db   2200 CTTGTCATGACTCAGACAGCAGGAAATGAAATTGAATTAGGAAATGAAGCTTGTGAATCT 2259

Qy     37 ------------------MetThrTyrSerThrGluArgLysThrAsnSerLeu------ 48
Db   2260 AAAGAUGGAATTGATAATGTGGACATGTCTTCAAAAGATTCRGGATCCAGTCTAATAAGA 2319

Qy     49 ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr 68
Db   2320 AGAAGATCAACTCGCAAAAGCATCCGTGGGCCACATGATCAAGACGGGGAACTTAGCACC 2379

Qy     69 ---GlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db   2380 AAAGAGGCTCTGGATGACGACGTACCTCCAGCTTCCTTTTGGCGGATCCTGAAGTTGAAT 2439

Qy     88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db   2440 TCAACTGAATGGCCTTATTTGTGGTTGGTGTATTTGTGCCATAATAAATGGAGGCTTG 2499

Qy    108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 127
Db   2500 CAACCAGCATTCTCCATAATATTTTCAAAGGTTGTAGGGGTTTTACAAAAAATGACACC 2559

Qy    128 ThrThrLeu---LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db   2560 CCTGAAATCCAGCGGCAGAACAGCAACTTGTTTTCTTTATTGTTCCTGATCCTTGGGATC 2619

Qy    147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db   2620 ATCTCTTTCATTACGTTTTTCCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTC 2679

Qy    167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
Db   2680 ACCAAGCGACTCCGATACATGGTCTTCAAATCCATGCTGAGACAGGACATAAGCTGGTTT 2739

Qy    187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db   2740 GATGACCCTAAAAACACCACAGGAGCGCTGACCACCAGGCTTGCCAATGACGCTGCTCAA 2799

Qy    207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db   2800 GTGAAAGGGGCTACAGGGTCTAGGCTTGCTGTTATTACCCAGAACATAGCAAATCTTGGG 2859

Qy    227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
Db   2860 ACAGGCATCATCATATCCCTGATCTACGGCTGGCAATTGACACTTTTACTCCTAGCAATT 2919

Qy    247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db   2920 GTTCCCATCATTGCTATAGCAGGAGTGGTTGAAATGAAAATGTTGTCTGGACAAGCGCTG 2979

Qy    267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db   2980 AAAGATAAGAAGGAACTAGAAGGTTCTGGGAAGATCGCTACAGAAGCAATTGAAAACTTT 3039

Qy    287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db   3040 CGCACTGTCGTCTCTTTGACTCGGGAGCAGAAGTTTGAAACTATGTATGCCCAGAGCTTG 3099

Qy    307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db   3100 CAGATACCATACAGAAATGCTTTGAAGAAAGCGCACGTCTTTGGGATCACTTTCTCCTTC 3159

Qy    327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
Db   3160 ACCCAGGCCATGATGTATTTCTCCTATGCTGCTTGTTTCCGGTTTGATGCCTACTTGGTG 3219

Qy    347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db   3220 GCACGAGAACTCATGACATTTGAAAATGTTCTGTTAGTATTCTCAGCTATTGTCTTTGGT 3279

Qy    367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db   3280 GCCATGGCAGTGGGGCAGGTCAGTTCATTCGCTCCTGACTACGCGAAAGCCAAAGTCTCG 3339

Qy    387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Db   3340 GCATCCCACATCATCAGGATCATTGAGAAAATCCCTGAGATTGACAGCTACAGCACGGAG 3399

Qy    407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
Db   3400 GGCTTGAAGCCTAATATGTTGGAAGGAAATGTGAAATTTAATGGAGTCATGTTCAACTAT 3459

Qy    427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db   3460 CCCACCCGACCCAACATCCCAGTGCTTCAGGGGCTGAGCCTAGAGGTGAAGAAAGGGCAG 3519

Qy    447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
Db   3520 ACGCTGGCCCTCGTGGGCAGCAGTGGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTTGAG 3579

Qy    467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db   3580 CGCTTCTATGACCCCATGGCCGGAACAGTGTTTCTAGATGGCAAAGAAATAAAGCAACTC 3639

Qy    487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db   3640 AATGTCCAGTGGCTCCGCGCCCACCTGGGCATTGTGTCCCAGGAGCCCATCCTGTTTGAC 3699

Qy    507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db   3700 TGCAGCATCGCCGAGAACATTGCCTACGGAGACAACAGCCGTGTCGTGTCTCATAAGGAG 3759

Qy    527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db   3760 ATCGTGAAGGCAGCCAAGGAGGCCAACATCCACCAGTTCATCGACTCACTGCCTGAGAAA 3819

Qy    547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Db   3820 TACAACACCAGAGTGGGAGACAAAGGGACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATC 3879

Qy    567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Db   3880 GCCATCGCGCGCGCCCTCGTCAGACAGCCTCACATCTTACTTCTGGATGAAGCGACATCA 3939

Qy    587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
Db   3940 GCTCTGGATACGGAGAGTGAAAAGGTCGTCCAGGAAGCGCTGGACAAAGCCAGGAAGGC 3999

Qy    607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db   4000 CGCACCTGCATTGTGATCGCGCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTG 4059

Qy    627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
Db   4060 GTGATTCAGAACGGCCAGGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCCCAGAAA 4119

Qy    647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db   4120 GGCATCTATTTCTCGATGGTCAGTGTGCAGGCT 4152

RESULT 8
```

```
AAF86127
ID   AAF86127 standard; cDNA; 4186 BP.
XX
AC   AAF86127;
XX
DT   25-JUN-2001  (first entry)
XX
DE   Cynomologous monkey P-glycoprotein cDNA.
XX
KW   Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   efflux pump; ss.
XX
OS   Macaca fascicularis.
XX
FH   Key             Location/Qualifiers
FT   CDS             100..3942
FT                   /*tag=  a
FT                   /product= "PGP"
FT                   /note= "P-glycoprotein"
XX
PN   WO200123565-A1.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26592.
XX
PR   28-SEP-1999;   99US-0156921.
PR   12-OCT-1999;   99US-0158818.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-Crespi DT,  Crespi CL;
XX
DR   WPI; 2001-316136/33.
DR   P-PSDB; AAB81064.
XX
PT   Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein
PT   (PGP) and homologous PGP polypeptides are useful for predicting
PT   bioavailability of compound and increasing PGP transporter activity in
PT   cell  -
XX
PS   Example 1; Page 51-57; 84pp; English.
XX
CC   This invention relates to a polynucleotide sequence encoding a
CC   cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the
CC   PGP protein. PGP, also known as multidrug transporter, MDR1 is a member
CC   of the ABC transporter superfamily. The enzyme serves as an efflux pump
CC   exporting small molecules across the cell membrane. The invention
CC   includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence
CC   and protein, and also that of an allelic variant. The PGP polynucleotide
CC   sequence is useful for increasing PGP transporter activity in a cell.
CC   Antisense sequences of the cDNA are useful for inhibiting PGP transport
CC   activity in a mammalian cell. They may also be used for increasing the
CC   bioavailability of a drug. The present sequence represents cDNA encoding
CC   cynomologous monkey P-glycoprotein.
XX
SQ   Sequence 4186 BP; 1226 A; 799 C; 1039 G; 1122 T; 0 other;

Alignment Scores:
Pred. No.:               1.17e-181     Length:         4186
Score:                   1873.00       Matches:        354
Percent Similarity:      74.29%        Conservative:   146
Best Local Similarity:   52.60%        Mismatches:     155
Query Match:             56.36%        Indels:         18
DB:                      22            Gaps:           5

US-09-873-409-1 (1-659) x AAF86127 (1-4186)

Qy        1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db     1912 GTCATTGTGGAGAAAGGAAATCATGATGAGCTCATGAAAGAGAAAGGCATTTACTTCAAA 1971

Qy       21 LeuValMetSerGlnAspIle---------------------LysLysAlaAspGlu--- 32
Db     1972 CTTGTCACAATGCAGACAGCAGGAAATGAAATTGAATTAGAAAATGCAGCTGATGAATCC 2031

Qy       33 ------GlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu------ 48
Db     2032 AAAAGTGAAATTGATACCTTGGAAATGTCTTCACATGATTCAGGATCCAGTCTAATAAGA 2091

Qy       49 ---------ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGlu 65
Db     2092 AAAAGATCCACTCGTAGGAGTGTCCGTGGATCACAAGGCCAA------GACAGAAAGCTT 2145

Qy       66 GluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLys 85
Db     2146 AGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAG 2205

Qy       86 LeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGly 105
Db     2206 CTAAATTTAACTGAGTGGCCTTATTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGA 2265

Qy      106 ThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsn 124
Db     2266 GGTCTGCAACCAGCATTTGCAGTAATATTTTCAAAGATTATAGGGATTTTTACAAGAAAT 2325

Qy      125 AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeu 144
Db     2326 GATGATGCCGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGTCCTT 2385

Qy      145 GlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGlu 164
Db     2386 GGAATTGTTTCTTTTATTACATTTTTCCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAG 2445

Qy      165 IleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAla 184
Db     2446 ATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGC 2505

Qy      185 TrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIle 204
Db     2506 TGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCT 2565

Qy      205 AlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsn 224
Db     2566 GCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTATAATTACCCAGAATATAGCAAAT 2625

Qy      225 MetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeu 244
Db     2626 CTTGGGACAGGAATAATTATATCCTTAATCTATGGTTGGCAACTGACACTGTTACTCTTA 2685

Qy      245 SerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPhe 264
Db     2686 GCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAA 2745

Qy      265 AlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGlu 284
Db     2746 GCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAA 2805

Qy      285 AsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGlu 304
Db     2806 AACTTCCGAACTGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGATCAG 2865

Qy      305 MetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyr 324
Db     2866 AGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATCACGTTT 2925

Qy      325 AlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyr 344
Db     2926 TCCTTCACGCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTAC 2985

Qy      345 LeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAla 364
Db     2986 TTGGTGGCACATAGTCTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTC 3045

Qy      365 TyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLys 384
```

```
Db   3046 TTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAA 3105

Qy    385 SerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSer 404
                |||||||||:::      ::::::||||||   |||   |||||||||   |||
Db   3106 GTATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGC 3165

Qy    405 GlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhe 424
             ||||||   ||||||:::|||   |||||||||:::   |||   ||||||   |||
Db   3166 ACAGAAGGCCTAAAGCCGAACACATTGGAAGGAAATGTCACATTTAATGAAGTTGTATTC 3225

Qy    425 PheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArg 444
             ||||||   |||   |||:::   :::|||:::|||||||||   :::::::::
Db   3226 AACTATCCCACCCGACTGGACATCCCAGTGCTTCAGGGGCTGAGCCTGGAAGTGAAGAAG 3285

Qy    445 GlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeu 464
          |||:::|||:::|||   |||||||||||||||||||||||||||   |||||||||
Db   3286 GGCCAGACGCTGGCCCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACGGTGGTCCAGCTC 3345

Qy    465 LeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLys 484
          |||:::|||   |||||||||:::   |||:::|||||||   ||||||   :::   |||
Db   3346 CTGGAGCGGTTCTATGACCCCTTGGCGGGGAAAGTGCTGCTTGACGGCAAAGAAATAAAG 3405

Qy    485 GluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeu 504
          :::||||||||||||||||||||||:::   :::   ||||||   |||||||||:::|||
Db   3406 CAACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTG 3465

Qy    505 PheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeu 524
          |||:::|||||||||:::|||||||||||||||||||||||||||||||||||||||||
Db   3466 TTTGACTGCAGCATTAGTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAG 3525

Qy    525 AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuPro 544
          :::||||||      ||||||      ||||||||||||:::|||||||||   ||||||
Db   3526 GAAGAGATCGTGAGGGCAGCCAAGGAGGCCAATATACACGCCTTCATCGAGTCACTGCCT 3585

Qy    545 GluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGln 564
             ||||||:::|||:::||||||   ||||||   ||||||||||||||||||||||
Db   3586 AATAAATATAGCACCAGAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAA 3645

Qy    565 ArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAla 584
          |||:::||||||||||||||||||:::::::::|||   -||||||||||||||||||
Db   3646 CGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCC 3705

Qy    585 ThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArg 604
          |||||||||||||||      :::|||||||||||||||   ||||||||||||||||||
Db   3706 ACATCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGA 3765

Qy    605 ThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeu 624
             ||||||||||||:::|||:::   ||||||||||||   |||||||||||||||||
Db   3766 GAAGGCCGTACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTA 3825

Qy    625 IleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArg 644
          |||||||||      ||||||:::::::||||||   ||||||||||||:::||||||
Db   3826 ATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCACGGCACACATCAGCAGCTGCTGGCA 3885

Qy    645 AsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 657
             :::   |||||||||   :::|||:::   |||:::
Db   3886 CAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3924

RESULT 9
AAF86128
ID   AAF86128 standard; cDNA; 4195 BP.
XX
AC   AAF86128;
XX
DT   25-JUN-2001  (first entry)
XX
DE   Cynomologous monkey P-glycoprotein cDNA variant 1.
XX
KW   Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   efflux pump; ss.
XX
OS   Macaca fascicularis.
XX
FH   Key             Location/Qualifiers
FT   CDS             100..3951
FT                   /*tag= a
FT                   /product= "PGP"
FT                   /note= "P-glycoprotein"
FT   misc_feature    376..384
FT                   /*tag= b
FT                   /note= "Insertion of 9 nucleotides relative to PGP
FT                          allelic variant AAF86127"
XX
PN   WO200123565-A1.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26592.
XX
PR   28-SEP-1999;   99US-0156921.
PR   12-OCT-1999;   99US-0158818.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-Crespi DT,  Crespi CL;
XX
DR   WPI; 2001-316136/33.
DR   P-PSDB; AAB81065.
XX
PT   Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein
PT   (PGP) and homologous PGP polypeptides are useful for predicting
PT   bioavailability of compound and increasing PGP transporter activity in
PT   cell  -
XX
PS   Example 1; Page 59-65; 84pp; English.
XX
CC   This invention relates to a polynucleotide sequence encoding a
CC   cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the
CC   PGP protein. PGP, also known as multidrug transporter, MDR1 is a member
CC   of the ABC transporter superfamily. The enzyme serves as an efflux pump
CC   exporting small molecules across the cell membrane. The invention
CC   includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence
CC   and protein, and also that of an allelic variant. The PGP polynucleotide
CC   sequence is useful for increasing PGP transporter activity in a cell.
CC   Antisense sequences of the cDNA are useful for inhibiting PGP transport
CC   activity in a mammalian cell. They may also be used for increasing the
CC   bioavailability of a drug. The present sequence represents cDNA encoding
CC   cynomologous monkey P-glycoprotein. This sequence contains a 9 nucleotide
CC   insert compared to the PGP allelic variant given in AAF86127.
XX
SQ   Sequence 4195 BP; 1230 A; 801 C; 1039 G; 1125 T; 0 other;

Alignment Scores:
Pred. No.:              1.18e-181       Length:         4195
Score:                  1873.00         Matches:        354
Percent Similarity:     74.29%          Conservative:   146
Best Local Similarity:  52.60%          Mismatches:     155
Query Match:            56.36%          Indels:         18
DB:                     22              Gaps:           5

US-09-873-409-1 (1-659) x AAF86128 (1-4195)

Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
          ::::::   |||||||||   |||   |||||||||   ::::::|||:::|||:::
Db   1921 GTCATTGTGGAGAAAGGAAATCATGATGAGCTCATGAAAGAGAAAGGCATTTACTTCAAA 1980

Qy     21 LeuValMetSerGlnAspIle---------------------LysLysAlaAspGlu--- 32
           ||||||      |||                               |||||||||
Db   1981 CTTGTCACAATGCAGACAGCAGGAAATGAAATTGAATTAGAAAATGCAGCTGATGAATCC 2040

Qy     33 ------GlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu------ 48
                :::::::::::::::      |||:::         :::||||||
Db   2041 AAAAGTGAAATTGATACCTTGGAAATGTCTTCACATGATTCAGGATCCAGTCTAATAAGA 2100

Qy     49 ---------ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGlu 65
                   ||||||:::      :::         |||:::
```

```
Db   2101 AAAAGATCCACTCGTAGGAGTGTCCGTGGATCACAAGGCCAA------GACAGAAAGCTT 2154

Qy     66 GluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLys 85
Db   2155 AGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAG 2214

Qy     86 LeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGly 105
Db   2215 CTAAATTTAACTGAGTGGCCTTATTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGA 2274

Qy    106 ThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsn 124
Db   2275 GGTCTGCAACCAGCATTTGCAGTAATATTTCAAAGATTATAGGGATTTTTACAAGAAAT 2334

Qy    125 AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeu 144
Db   2335 GATGATGCCGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGTCCTT 2394

Qy    145 GlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGlu 164
Db   2395 GGAATTGTTTCTTTTATTACATTTTTCCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAG 2454

Qy    165 IleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAla 184
Db   2455 ATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGC 2514

Qy    185 TrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIle 204
Db   2515 TGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCT 2574

Qy    205 AlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsn 224
Db   2575 GCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTATAATTACCCAGAATATAGCAAAT 2634

Qy    225 MetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeu 244
Db   2635 CTTGGGACAGGAATAATTATATCCTTAATCTATGGTTGGCAACTGACACTGTTACTCTTA 2694

Qy    245 SerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPhe 264
Db   2695 GCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAA 2754

Qy    265 AlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGlu 284
Db   2755 GCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAA 2814

Qy    285 AsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGlu 304
Db   2815 AACTTCCGAACTGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGATCAG 2874

Qy    305 MetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyr 324
Db   2875 AGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATCACGTTT 2934

Qy    325 AlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyr 344
Db   2935 TCCTTCACGCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTAC 2994

Qy    345 LeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAla 364
Db   2995 TTGGTGGCACATAGTCTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTC 3054

Qy    365 TyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLys 384
Db   3055 TTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAA 3114

Qy    385 SerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSer 404
Db   3115 GTATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGC 3174

Qy    405 GlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhe 424
Db   3175 ACAGAAGGCCTAAAGCCGAACACATTGGAAGGAAATGTCACATTTAATGAAGTTGTATTC 3234

Qy    425 PheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArg 444
Db   3235 AACTATCCCACCCGACTGGACATCCCAGTGCTTCAGGGGCTGAGCCTGGAAGTGAAGAAG 3294

Qy    445 GlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeu 464
Db   3295 GGCCAGACGCTGGCCCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACGGTGGTCCAGCTC 3354

Qy    465 LeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLys 484
Db   3355 CTGGAGCGGTTCTATGACCCCTTGGCGGGGAAAGTGCTGCTTGACGGCAAAGAAATAAAG 3414

Qy    485 GluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeu 504
Db   3415 CAACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTG 3474

Qy    505 PheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeu 524
Db   3475 TTTGACTGCAGCATTAGTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAG 3534

Qy    525 AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuPro 544
Db   3535 GAAGAGATCGTGAGGGCAGCCAAGGAGGCCAATATACACGCCTTCATCGAGTCACTGCCT 3594

Qy    545 GluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGln 564
Db   3595 AATAAATATAGCACCAGAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAA 3654

Qy    565 ArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAla 584
Db   3655 CGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCC 3714

Qy    585 ThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArg 604
Db   3715 ACATCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGA 3774

Qy    605 ThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeu 624
Db   3775 GAAGGCCGTACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTA 3834

Qy    625 IleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArg 644
Db   3835 ATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCACGGCACACATCAGCAGCTGCTGGCA 3894

Qy    645 AsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db   3895 CAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3933
```

RESULT 10
AAQ70907

```
ID   AAQ70907 standard; DNA; 2726 BP.
XX
AC   AAQ70907;
XX
DT   27-MAR-1995  (first entry)
XX
DE   Multidrug-resistance gene MDR-1.
XX
KW   Multidrug-resistance; MDR-1 gene; ds.
XX
OS   Synthetic.
XX
PN   WO9417206-A.
XX
PD   04-AUG-1994.
XX
PF   12-NOV-1993;   93WO-US10883.
XX
PR   27-JAN-1993;   93US-0010433.
PR   16-DEC-1993;   93US-0168621.
XX
PA   (ONCO-) ONCOR INC.
```

```
XX
PI   Bhatnagar SK,  George AL;
XX
DR   WPI; 1994-264118/32.
XX
PT   Enzymatic amplification of target nucleic acid sequences in a
PT   mixt. - to detect a mutation or allele in the target, e.g. that
PT   causes a genetic disease
XX
PS   Disclosure; Page 30; 50pp; English.
XX
CC   This gene is used as an example of a target DNA in a method for the
CC   amplification of nucleic acid sequences in a mixture using various
CC   DNA probes/primers.  The method may be used to detect a mutation or
CC   allele in the target, e.g. that that causes a genetic disease.
XX
SQ   Sequence 2726 BP; 818 A; 538 C; 642 G; 728 T; 0 other;

Alignment Scores:
Pred. No.:              7.82e-182     Length:         2726
Score:                  1872.00       Matches:        356
Percent Similarity:     74.22%        Conservative:   142
Best Local Similarity:  53.06%        Mismatches:     159
Query Match:            56.33%        Indels:         14
DB:                     15            Gaps:           4

US-09-873-409-1 (1-659) x AAQ70907 (1-2726)

Qy       1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db     317 GTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAA 376

Qy      21 LeuValMetSerGln------------------AspIleLysLysAlaAspGluGlnMet 34
Db     377 CTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCC 436

Qy      35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
Db     437 AAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGA 496

Qy      54 ---------------LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
Db     497 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACC 556

Qy      68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db     557 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAAT 616

Qy      88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db     617 TTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTG 676

Qy     108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
Db     677 CAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGAT 736

Qy     127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db     737 CCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATT 796

Qy     147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db     797 ATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTC 856

Qy     167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
Db     857 ACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 916

Qy     187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db     917 GATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAA 976

Qy     207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db     977 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGG 1036

Qy     227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
Db    1037 ACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATT 1096

Qy     247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db    1097 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 1156

Qy     267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db    1157 AAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 1216

Qy     287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db    1217 CGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTG 1276

Qy     307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db    1277 CAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTC 1336

Qy     327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
Db    1337 ACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTG 1396

Qy     347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db    1397 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGT 1456

Qy     367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db    1457 GCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCA 1516

Qy     387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Db    1517 GCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAA 1576

Qy     407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
Db    1577 GGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTAT 1636

Qy     427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db    1637 CCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAG 1696

Qy     447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
Db    1697 ACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAG 1756

Qy     467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db    1757 CGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTG 1816

Qy     487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db    1817 AATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGAC 1876

Qy     507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db    1877 TGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAG 1936

Qy     527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db    1937 ATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAA 1996

Qy     547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Db    1997 TATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATT 2056

Qy     567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
```

```
Db   2057 GCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTGCTTTTGGATGAAGCCACGTCA 2116

Qy    587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
          ||||||||    :::||||||||||||||||||   ||||||||||||||||    |||
Db   2117 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 2176

Qy    607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
          ||||||||:::|||:::   |||||||||||   |||||||||||||||||||||||
Db   2177 CGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTG 2236

Qy    627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
          |||         ||||||:::::||||||   ||||||||||||:::||||||      :::
Db   2237 GTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAA 2296

Qy    647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
             |||||||||   :::|||:::   |||:::
Db   2297 GGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 2329
```

RESULT 11
AAQ70916
ID AAQ70916 standard; DNA; 2726 BP.
XX
AC AAQ70916;
XX
DT 27-MAR-1995 (first entry)
XX
DE Multidrug-resistance gene MDR-1.
XX
KW Multidrug-resistance; MDR-1 gene; ds.
XX
OS Synthetic.
XX
PN WO9417210-A.
XX
PD 04-AUG-1994.
XX
PF 25-JAN-1994; 94WO-US00748.
XX
PR 27-JAN-1993; 93US-0010433.
PR 16-DEC-1993; 93US-0168621.
XX
PA (ONCO-) ONCOR INC.
XX
PI Bhatnagar SK, George AL;
XX
DR WPI; 1994-264122/32.
XX
PT Enzymatic amplification of target nucleic acid sequences to form
PT exact or modified copies - has increased fidelity and can identify
PT point mutations or allele(s)
XX
PS Disclosure; Page 30; 69pp; English.
XX
CC This gene is used as an example of a target DNA in a method for the
CC amplification of nucleic acid sequences in a mixture using various
CC DNA probes/primers. The method may be used to detect a mutation or
CC allele in the target, e.g. that that causes a genetic disease.
XX
SQ Sequence 2726 BP; 818 A; 538 C; 642 G; 728 T; 0 other;

Alignment Scores:

Pred. No.:	7.82e-182	Length:	2726
Score:	1872.00	Matches:	356
Percent Similarity:	74.22%	Conservative:	142
Best Local Similarity:	53.06%	Mismatches:	159
Query Match:	56.33%	Indels:	14
DB:	15	Gaps:	4

US-09-873-409-1 (1-659) x AAQ70916 (1-2726)

```
Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
          ::::::   |||||||||   |||   |||||||||   ::::::|||:::|||:::
Db    317 GTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAA 376

Qy     21 LeuValMetSerGln------------------AspIleLysLysAlaAspGluGlnMet 34
          |||||      |||                   :::::::::   |||   ::::::
Db    377 CTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCC 436

Qy     35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
          :::|||         :::   |||         :::|||                  :::
Db    437 AAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGA 496

Qy     54 ---------------LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
                         :::|||::::::   |||         |||:::            :::
Db    497 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACC 556

Qy     68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
          ::::::   :::      |||:::|||   ||||||         :::|||:::|||||||||
Db    557 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAAT 616

Qy     88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
                |||||||||:::   |||:::|||         :::::::::||||||   :::
Db    617 TTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTG 676

Qy    108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
             |||   |||:::|||||||||:::|||||||||      :::|||         :::|||
Db    677 CAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGAT 736

Qy    127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
                |||   :::   ::::::   ::::::|||::::::|||:::   ||||||:::
Db    737 CCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATT 796

Qy    147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
          |||   |||:::::::::|||:::|||||||:::||||||||||||||||||
Db    797 ATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTC 856

Qy    167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
          |||   ||||||||||::::::   |||::::::||||||   ||||||::::::||||||
Db    857 ACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 916

Qy    187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
          |||:::   :::|||:::||||||   |||||||||   ||||||   |||   ||||||
Db    917 GATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAA 976

Qy    207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
          ::::::||||||   |||||||||:::   |||:::|||||||||      |||:::|||
Db    977 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGG 1036

Qy    227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
             :::|||||||||||||||||||||||||||   |||:::|||:::|||
Db   1037 ACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATT 1096

Qy    247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
             |||::::::|||:::   |||::::::|||         ::::::|||   |||
Db   1097 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 1156

Qy    267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
          |||||||||:::||||||:::   ||||||||||||||||||||||||:::||||||
Db   1157 AAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 1216

Qy    287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
          ||||||:::|||||||||||||:::|||:::   ||||||   ||||||   :::   |||
Db   1217 CGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTG 1276

Qy    307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
          |||      :::|||||||:::   ||||||   |||   |||      ::::::|||
Db   1277 CAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTC 1336

Qy    327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
          :::   |||   :::||||||:::||||||      ||||||||||||||||||:::
Db   1337 ACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTG 1396

Qy    347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
                   |||:::   |||   :::   :::||||||:::|||:::   :::|||
Db   1397 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGT 1456
```

```
Qy     367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
           ||||||||:::|||:::               ||||||:::|||:::||||||||
Db    1457 GCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCA 1516

Qy     387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
           |||||||||:::         ::::::||||||   |||   |||||||||   |||   |||
Db    1517 GCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAA 1576

Qy     407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
           |||      |||:::|||   |||||||||:::   |||. ||||||   |||   |||
Db    1577 GGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTAT 1636

Qy     427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
           |||   |||||||||:::   :::|||:::||||||||||||   :::::::::|||:::
Db    1637 CCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAG 1696

Qy     447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
           |||:::|||   |||||||||||||||||||||||||||   ||||||||||||:::
Db    1697 ACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAG 1756

Qy     467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
           |||   |||||||||:::   |||:::||||||   |||||   :::   |||   |||
Db    1757 CGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTG 1816

Qy     487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
           ||||||||||||||||||:::   :::   ||||||   |||||||||:::||||||:::
Db    1817 AATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGAC 1876

Qy     507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
           |||||||||||||||||||||||||||||||||||||||||||||||||       :::|||
Db    1877 TGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAG 1936

Qy     527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
           |||         ||||||         ||||||||||||:::|||||||||   ||||||   |||
Db    1937 ATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAA 1996

Qy     547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
           |||:::|||:::||||||      ||||||   ||||||||||||||||||||||||:::
Db    1997 TATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATT 2056

Qy     567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
           ||||||||||||||||||:::::::|||   ||||||||||||||||||||||||||||
Db    2057 GCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCA 2116

Qy     587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
           ||||||||   :::||||||||||||||||||   ||||||||||||||||||   |||
Db    2117 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 2176

Qy     607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
           |||||||||:::|||:::   ||||||||||||   ||||||||||||||||||||||||
Db    2177 CGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTG 2236

Qy     627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
           |||         ||||||::::::||||||   ||||||||||||:::||||||      :::
Db    2237 GTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAA 2296

Qy     647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
                 |||||||||   :::|||:::   |||:::
Db    2297 GGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 2329
```

RESULT 12
AAT43322
ID   AAT43322 standard; DNA; 2726 BP.
XX
AC   AAT43322;
XX
DT   01-SEP-1997 (first entry)
XX
DE   Multidrug resistance gene-1.
XX
KW   PCR; primer; amplify; polymerase chain reaction; ligase chain reaction;
KW   LCR; human; multidrug resistance gene; MDR-1; ds.
XX
OS   Synthetic.
XX
PN   WO9639537-A1.
XX
PD   12-DEC-1996.
XX
PF   04-JUN-1996;   96WO-US08841.
XX
PR   05-JUN-1995;   95US-0461823.
XX
PA   (ONCO-) ONCOR INC.
XX
PI   Bhatnagar SK, George AL, Nazarenko I;
XX
DR   WPI; 1997-043158/04.
XX
PT   Amplification method avoiding strand displacement by polymerase -
PT   used in the detection of mutation(s) and allele(s) associated with
PT   genetic disease and cancer
XX
PS   Example 1; Page 50-51; 92pp; English.
XX
CC   This sequence represents the multidrug resistance gene (MDR-1) amplified
CC   by the primers shown in AAT43320 and AAT43321. This sequence can be used
CC   as a target in the method of the invention, for enzymatically amplifying
CC   a target nucleic acid (TNA) sequence contained in a nucleic acid or
CC   mixture of nucleic acids while avoiding strand displacement by
CC   polymerase. The method comprises using three primers, one complementary
CC   to a first segment of the TNA, a second complementary to a second segment
CC   of the TNA, which is adjacent to the first primer, and a third which is
CC   similar to the first segment of the TNA, and is complementary to a
CC   portion of the first primer. The first two primers are hybridised to the
CC   TNA, and a fused amplification product is created from the TNA using
CC   these two primers. The fused product is dissociated, and hybridised to
CC   the third primer, which is then extended. The extended modified
CC   amplification product is hybridised to the first two primers, and these
CC   two primers are ligated. Each of the primers may be labelled using a
CC   different label so that the method can be used to detect the presence of
CC   a mutation or allele by detecting whether the labelled primer is
CC   contained within the fused amplification product or the extended
CC   amplification product. The method of the invention combines certain
CC   aspects of ligase chain reaction (LCR) and polymerase chain reaction
CC   (PCR), but is improved compared to LCR due to the reduced number of
CC   primers needed, and the fact that the entire TNA sequence does not need
CC   to be known.
XX
SQ   Sequence 2726 BP; 818 A; 538 C; 642 G; 728 T; 0 other;

```
Alignment Scores:
Pred. No.:              7.82e-182      Length:         2726
Score:                  1872.00        Matches:        356
Percent Similarity:     74.22%         Conservative:   142
Best Local Similarity:  53.06%         Mismatches:     159
Query Match:            56.33%         Indels:         14
DB:                     18             Gaps:           4
```

US-09-873-409-1 (1-659) x AAT43322 (1-2726)

```
Qy       1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
              ::::::   |||||||||   |||   |||||||||   ::::::|||:::|||:::
Db     317 GTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAA 376

Qy      21 LeuValMetSerGln------------------AspIleLysLysAlaAspGluGlnMet 34
           ||||||      |||                  :::::::::   |||   ::::::
Db     377 CTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCC 436

Qy      35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
              :::|||         :::   |||      :::|||               :::
Db     437 AAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGA 496

Qy      54 ---------------LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
                          :::|||::::::   |||         |||:::         :::
Db     497 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACC 556
```

```
Qy       68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
               ::::::   :::   |||:::|||   ||||||       :::|||:::||||||||
Db      557 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAAT 616

Qy       88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
                     |||||||||:::    |||:::|||       :::::::::|||||    :::
Db      617 TTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTG 676

Qy      108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
               |||    |||:::|||||||||:::||||||||       :::|||         :::|||
Db      677 CAACCAGCATTTGCAATAATATTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGAT 736

Qy      127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
                  |||    :::    ::::::    ::::::|||::::::|||:::     ||||||:::
Db      737 CCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATT 796

Qy      147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
            |||    |||:::::::::|||:::||||||       :::|||:::|||||||||||||||
Db      797 ATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTC 856

Qy      167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
            |||    |||||||||::::::    |||::::::||||||    ||||||::::::||||||
Db      857 ACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 916

Qy      187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
            |||:::    :::|||:::||||||   |||||||||    ||||||    |||    ||||||
Db      917 GATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAA 976

Qy      207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
            ::::::||||||    |||||||||:::          |||:::|||          |||:::|||
Db      977 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGG 1036

Qy      227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
               :::||||||||||||||||||||||||:::::::|||    |||:::|||:::|||
Db     1037 ACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATT 1096

Qy      247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
               |||:::::::|||:::    |||:::::::|||       ::::::|||    |||
Db     1097 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 1156

Qy      267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
            |||||||||:::||||||:::    |||||||||||||||||||||||:::||||||
Db     1157 AAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 1216

Qy      287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
            ||||||:::|||||||||||||:::|||:::    ||||||    ||||||    :::    |||
Db     1217 CGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTG 1276

Qy      307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
            |||       :::||||||:::    :::||||||   |||   |||       ::::::|||
Db     1277 CAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTC 1336

Qy      327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
            :::    |||    :::||||||:::||||||       |||||||||||||||||||:::
Db     1337 ACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTG 1396

Qy      347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
                     |||:::    |||    :::    :::||||||:::|||:::    :::|||
Db     1397 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTCAGCTGTTGTCTTTGGT 1456

Qy      367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
            |||||||||:::|||:::             ||||||:::|||:::|||||||||
Db     1457 GCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCA 1516

Qy      387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
            |||||||||:::          ::::::||||||   |||   |||||||||   |||   |||
Db     1517 GCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAA 1576

Qy      407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
            |||          |||:::|||   |||||||||:::   |||   ||||||   |||   |||
Db     1577 GGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTAT 1636

Qy      427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
            |||   |||||||||:::    :::|||:::|||||||||||||    :::::::::|||:::
Db     1637 CCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAG 1696

Qy      447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
            |||:::|||   ||||||||||||||||||||||||||||    ||||||||||||:::
Db     1697 ACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAG 1756

Qy      467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
            |||    |||||||||:::    |||:::||||||    ||||||    :::    |||    |||
Db     1757 CGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTG 1816

Qy      487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
            ||||||||||||||||:::    :::    ||||||    |||||||||:::||||||:::
Db     1817 AATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGAC 1876

Qy      507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
            ||||||||||||||||||||||||||||||||||||||||||||||          :::|||
Db     1877 TGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAG 1936

Qy      527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
            |||       ||||||       ||||||||||||:::||||||||||    ||||||    |||
Db     1937 ATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAA 1996

Qy      547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
            |||:::|||:::||||||    ||||||    |||||||||||||||||||||||||||:::
Db     1997 TATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATT 2056

Qy      567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
            ||||||||||||||||||:::::::::|||    ||||||||||||||||||||||||
Db     2057 GCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCA 2116

Qy      587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
            |||||||||    :::||||||||||||||||||    ||||||||||||||||||    |||
Db     2117 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 2176

Qy      607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
            |||||||||:::|||:::    ||||||||||||||||||||||||||||
Db     2177 CGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTG 2236

Qy      627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
            |||    ||||||::::::||||||    ||||||||||||:::||||||          :::
Db     2237 GTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAA 2296

Qy      647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
               ||||||||    :::|||:::   |||:::
Db     2297 GGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 2329
```

RESULT 13
ABL91687
ID   ABL91687 standard; DNA; 3840 BP.
XX
AC   ABL91687;
XX
DT   28-MAY-2002 (first entry)
XX
DE   Human polynucleotide SEQ ID NO 30.
XX
KW   Human; HIV; HCV; gene expression; oligoribonucleotide; tumour; pathogen;
KW   Plasmodium; virus; viroid; cytokine; prion; antisense oligonucleotide;
KW   cytostatic; virucide; protozoacide; antibacterial; ds.
XX
OS   Homo sapiens.
XX
PN   DE10100586-C1.
XX
PD   11-APR-2002.
XX
PF   09-JAN-2001; 2001DE-1000586.
XX
PR   09-JAN-2001; 2001DE-1000586.
XX
PA   (RIBO-) RIBOPHARMA AG.

```
XX
PI   Kreutzer R,  Limmer S,  Rost S,  Hadwiger P;
XX
DR   WPI; 2002-270454/32.
XX
PT   Inhibiting gene expression in cells, useful for e.g. treating tumors,
PT   by introducing double-stranded complementary oligoRNA having unpaired
PT   terminal bases  -
XX
PS   Claim 13; Page 28-30; 104pp; German.
XX
CC   The invention relates to a method for inhibiting expression of a target
CC   gene (ABL91658-ABL91797) in a cell by introducing at least one
CC   oligoribonucleotide that has a double-stranded structure consisting of at
CC   most 49 sequential nucleotide pairs, with at least part of one strand
CC   complementary with the target gene and has at least one end a
CC   single-stranded segment of 1-4 nt. The method provides
CC   oligoribonucleotides for antisense inhibition of gene expression useful
CC   e.g. for treating tumours but the oligoribonucleotides may also be
CC   directed against genes present in pathogens (e.g. Plasmodium or
CC   viruses/viroids, pathogenic on humans, animals or plants) or against
CC   cytokine, Id, developmental or prion genes. The method provides more
CC   effective inhibition of gene expression than use of known
CC   oligonucleotides, probably because the unpaired overhang increases
CC   stability and thus intracellular concentration.
XX
SQ   Sequence 3840 BP; 1130 A; 738 C; 957 G; 1015 T; 0 other;

Alignment Scores:
Pred. No.:              1.31e-181     Length:        3840
Score:                  1872.00       Matches:       356
Percent Similarity:     74.22%        Conservative:  142
Best Local Similarity:  53.06%        Mismatches:    159
Query Match:            56.33%        Indels:        14
DB:                     24            Gaps:          4

US-09-873-409-1 (1-659) x ABL91687 (1-3840)

Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db   1810 GTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAA 1869

Qy     21 LeuValMetSerGln------------------AspIleLysLysAlaAspGluGlnMet 34
Db   1870 CTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCC 1929

Qy     35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
Db   1930 AAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGA 1989

Qy     54 ---------------LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
Db   1990 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACC 2049

Qy     68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db   2050 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAAT 2109

Qy     88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db   2110 TTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTG 2169

Qy    108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
Db   2170 CAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGAT 2229

Qy    127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db   2230 CCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATT 2289

Qy    147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db   2290 ATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTC 2349

Qy    167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
Db   2350 ACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 2409

Qy    187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db   2410 GATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAA 2469

Qy    207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db   2470 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGG 2529

Qy    227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
Db   2530 ACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATT 2589

Qy    247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db   2590 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 2649

Qy    267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db   2650 AAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 2709

Qy    287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db   2710 CGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTG 2769

Qy    307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db   2770 CAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTC 2829

Qy    327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
Db   2830 ACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTG 2889

Qy    347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db   2890 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGT 2949

Qy    367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db   2950 GCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCA 3009

Qy    387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Db   3010 GCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAA 3069

Qy    407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
Db   3070 GGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTAT 3129

Qy    427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db   3130 CCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAG 3189

Qy    447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
Db   3190 ACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAG 3249

Qy    467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db   3250 CGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTG 3309

Qy    487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db   3310 AATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGAC 3369

Qy    507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db   3370 TGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAG 3429
```

```
Qy    527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
          |||       ||||||      |||||||||||:::||||||||    ||||||    |||
Db   3430 ATTGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAA 3489

Qy    547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
          |||:::|||:::||||||   |||||   ||||||||||||||||||||||||||||:::
Db   3490 TATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATT 3549

Qy    567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
          ||||||||||||||||||:::::::::|||   ||||||||||||||||||||||||
Db   3550 GCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCA 3609

Qy    587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
          |||||||||   :::||||||||||||||||||   ||||||||||||||||||   |||
Db   3610 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 3669

Qy    607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
          |||||||||:::|||:::   ||||||||||||   ||||||||||||||||||||||||
Db   3670 CGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTG 3729

Qy    627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
          |||      ||||||::::::||||||   ||||||||||||:::||||||      :::
Db   3730 GTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAA 3789

Qy    647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
             |||||||||   :::|||:::   |||:::
Db   3790 GGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3822
```

RESULT 14
AAZ49332

```
ID   AAZ49332 standard; cDNA; 3860 BP.
XX
AC   AAZ49332;
XX
DT   14-MAR-2000  (first entry)
XX
DE   Human wild-type multidrug resistance-1 (MDR-1) cDNA.
XX
KW   Multidrug resistance; MDR-1; P-glycoprotein;
KW   transmembrane efflux pump; haematopoietic stem cell; transduction;
KW   bone marrow transplantation; chemotherapy; radiation therapy; cancer;
KW   gene therapy; gene replacement; genetic defect; thalassaemia;
KW   Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
KW   cytokine; wild-type; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3843
FT                   /*tag=  a
FT                   /product= "Human wild-type MDR-1 protein"
FT   mutation        replace (553..555, GTT)
FT                   /*tag=  b
FT                   /note= "cDNA sequence of G185V human mutant MDR-1 given
FT                          in AAZ49333"
XX
PN   WO9961589-A2.
XX
PD   02-DEC-1999.
XX
PF   27-MAY-1999;   99WO-US11825.
XX
PR   28-MAY-1998;   98US-0086988.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI   Sorrentino B,  Bunting K;
XX
DR   WPI; 2000-072615/06.
DR   P-PSDB; AAY58186.
XX
PT   Ex vivo expansion of hematopoietic stem cells transduced with a
PT   sequence encoding human multidrug resistance-1, used for bone marrow
PT   transplantation -
XX
PS   Claim 10; Page 68-70; 113pp; English.
XX
CC   This sequence represents cDNA encoding human wild-type
CC   multidrug resistance protein MDR-1. MDR-1 is a transmembrane
CC   efflux pump, responsible for the export of drugs from cells,
CC   particularly cancer cells. Wild-type MDR-1 shows increased
CC   resistance to etoposide and decreased resistance to vinca
CC   alkaloids compared with a mutant form (AAY58187) where the Gly at
CC   position 185 is replaced by Val. The invention relates to transducing
CC   haematopoietic stem cells with nucleic acid encoding an MDR protein
CC   and culturing the modified cells. The modified haematopoietic stem
CC   cells are useful in bone marrow transplantation (to reconstitute
CC   haematopoietic systems in patients who have undergone chemotherapy or
CC   radiation therapy) and in ex vivo gene therapy of genetic defects in
CC   cells derived from haematopoietic stem cells, e.g., thalassaemia,
CC   Gauchier's disease, sickle cell anaemia or leukaemia. The modified
CC   cells can also be used to identify factors involved in regulating
CC   proliferation and differentiation in haematopoietic stem cells.
CC   Haematopoietic stem cells that express MDR-1 will be protected against
CC   chemotherapeutic agents, so can be engrafted while the patient is
CC   undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells
CC   provides sufficient cells to permit standard biochemical analysis.
CC   Overexpression of MDR-1 allows cytokine-driven expansion of
CC   haematopoietic stem cells by at least 10-fold compared with a maximum
CC   of 4-fold in known procedures.
XX
SQ   Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;

Alignment Scores:
Pred. No.:              1.32e-181     Length:        3860
Score:                  1872.00       Matches:       356
Percent Similarity:     74.22%        Conservative:  142
Best Local Similarity:  53.06%        Mismatches:    159
Query Match:            56.33%        Indels:        14
DB:                     21            Gaps:          4

US-09-873-409-1 (1-659) x AAZ49332 (1-3860)

Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
            ::::::   |||||||||   |||   |||||||||   ::::::|||:::|||:::
Db   1813 GTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAA 1872

Qy     21 LeuValMetSerGln------------------AspIleLysLysAlaAspGluGlnMet 34
          ||||||   |||                     :::::::::   |||   ::::::
Db   1873 CTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCC 1932

Qy     35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
          :::|||      :::   |||      :::|||                     :::
Db   1933 AAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGA 1992

Qy     54 ---------------LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
                         :::|||::::::      |||         |||:::   :::
Db   1993 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACC 2052

Qy     68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
             ::::::   :::   |||:::|||   ||||||      :::|||:::|||||||||
Db   2053 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAAT 2112

Qy     88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
                |||||||||:::   |||:::|||      :::::::::||||||   :::
Db   2113 TTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTG 2172

Qy    108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
             |||   |||:::|||||||||:::|||||||||      :::|||         :::|||
Db   2173 CAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGAT 2232

Qy    127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
                |||   :::   ::::::   ::::::|||::::::|||:::   ||||||:::
Db   2233 CCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATT 2292

Qy    147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
```

```
Db   2293 ATTTCTTTTATTACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTC 2352

Qy    167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
Db   2353 ACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 2412

Qy    187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db   2413 GATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAA 2472

Qy    207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db   2473 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGG 2532

Qy    227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
Db   2533 ACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATT 2592

Qy    247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db   2593 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 2652

Qy    267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db   2653 AAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 2712

Qy    287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db   2713 CGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTG 2772

Qy    307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db   2773 CAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTC 2832

Qy    327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
Db   2833 ACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTG 2892

Qy    347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db   2893 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGT 2952

Qy    367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db   2953 GCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCA 3012

Qy    387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Db   3013 GCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAA 3072

Qy    407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
Db   3073 GGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTAT 3132

Qy    427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db   3133 CCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAG 3192

Qy    447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
Db   3193 ACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAG 3252

Qy    467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db   3253 CGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTG 3312

Qy    487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db   3313 AATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGAC 3372

Qy    507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db   3373 TGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAG 3432

Qy    527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db   3433 ATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAA 3492

Qy    547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Db   3493 TATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATT 3552

Qy    567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Db   3553 GCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCA 3612

Qy    587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
Db   3613 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 3672

Qy    607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db   3673 CGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTG 3732

Qy    627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
Db   3733 GTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAA 3792

Qy    647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db   3793 GGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3825
```

RESULT 15
AAZ49333

```
ID   AAZ49333 standard; cDNA; 3860 BP.
XX
AC   AAZ49333;
XX
DT   14-MAR-2000  (first entry)
XX
DE   Human G185V mutant multidrug resistance-1 (MDR-1) cDNA.
XX
KW   Multidrug resistance; MDR-1; P-glycoprotein;
KW   transmembrane efflux pump; haematopoietic stem cell; transduction;
KW   bone marrow transplantation; chemotherapy; radiation therapy; cancer;
KW   gene therapy; gene replacement; genetic defect; thalassaemia;
KW   Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
KW   cytokine; mutant; ds.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3843
FT                   /*tag=  a
FT                   /product= "Human G185V mutant MDR-1 protein"
XX
PN   WO9961589-A2.
XX
PD   02-DEC-1999.
XX
PF   27-MAY-1999;   99WO-US11825.
XX
PR   28-MAY-1998;   98US-0086988.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI   Sorrentino B,  Bunting K;
XX
DR   WPI; 2000-072615/06.
DR   P-PSDB; AAY58187.
XX
PT   Ex vivo expansion of hematopoietic stem cells transduced with a
PT   sequence encoding human multidrug resistance-1, used for bone marrow
PT   transplantation -
```

```
XX
PS   Example 1; Page 79-82; 113pp; English.
XX
CC   This sequence represents cDNA encoding human G185V mutant multidrug
CC   resistance protein MDR-1, where the the Gly residue at postion 185
CC   of the wild-type protein (AAY58186) is replaced by Val. MDR-1 is a
CC   transmembrane efflux pump, responsible for the export of drugs from
CC   cells, particularly cancer cells. The wild-type MDR-1 shows increased
CC   resistance to etoposide and decreased resistance to vinca alkaloids
CC   compared with the G185V mutant. The invention relates to transducing
CC   haematopoietic stem cells with nucleic acid encoding an MDR protein
CC   and culturing the modified cells. The modified haematopoietic stem
CC   cells are useful in bone marrow transplantation (to reconstitute
CC   haematopoietic systems in patients who have undergone chemotherapy or
CC   radiation therapy) and in ex vivo gene therapy of genetic defects in
CC   cells derived from haematopoietic stem cells, e.g., thalassaemia,
CC   Gauchier's disease, sickle cell anaemia or leukaemia. The modified
CC   cells can also be used to identify factors involved in regulating
CC   proliferation and differentiation in haematopoietic stem cells.
CC   Haematopoietic stem cells that express MDR-1 will be protected against
CC   chemotherapeutic agents, so can be engrafted while the patient is
CC   undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells
CC   provides sufficient cells to permit standard biochemical analysis.
CC   Overexpression of MDR-1 allows cytokine-driven expansion of
CC   haematopoietic stem cells by at least 10-fold compared with a maximum
CC   of 4-fold in known procedures.
XX
SQ   Sequence 3860 BP; 1135 A; 746 C; 956 G; 1023 T; 0 other;

Alignment Scores:
Pred. No.:              1.32e-181     Length:        3860
Score:                  1872.00       Matches:       356
Percent Similarity:     74.22%        Conservative:  142
Best Local Similarity:  53.06%        Mismatches:    159
Query Match:            56.33%        Indels:        14
DB:                     21            Gaps:          4

US-09-873-409-1 (1-659) x AAZ49333 (1-3860)

Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
          ::::::   ||||||||   |||   ||||||||    ::::::|||:::|||:::
Db   1813 GTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAA 1872

Qy     21 LeuValMetSerGln------------------AspIleLysLysAlaAspGluGlnMet 34
          ||||||   |||                     :::::::::   |||   ::::::
Db   1873 CTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCC 1932

Qy     35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
          :::|||         :::   |||      :::|||                  :::
Db   1933 AAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGA 1992

Qy     54 ---------------LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
                         :::|||::::::      |||      |||:::            :::
Db   1993 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACC 2052

Qy     68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
          ::::::   :::      |||:::|||   ||||||      :::|||:::||||||||
Db   2053 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAAT 2112

Qy     88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
                   ||||||||:::   |||:::|||      :::::::::||||||   :::
Db   2113 TTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTG 2172

Qy    108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
             |||   |||:::|||||||||:::||||||||      :::|||      :::|||
Db   2173 CAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGAT 2232

Qy    127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
                |||   :::   ::::::   ::::::|||::::::|||:::   ||||||:::
Db   2233 CCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATT 2292

Qy    147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
          |||   |||:::::::::|||:::||||||      :::|||:::||||||||||||||
Db   2293 ATTTCTTTTATTACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTC 2352

Qy    167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
          |||   ||||||||::::::   |||::::::||||||   ||||||::::::||||||
Db   2353 ACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 2412

Qy    187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
          |||:::   :::|||:::||||||   |||||||||   ||||||   |||   ||||||
Db   2413 GATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAA 2472

Qy    207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
          ::::::||||||   |||||||||:::   |||:::|||||||||      |||:::|||
Db   2473 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGG 2532

Qy    227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
                :::|||||||||||||||||||||||||::::::|||   |||:::|||:::|||
Db   2533 ACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATT 2592

Qy    247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
             |||::::::|||:::   |||::::::|||         ::::::|||   |||
Db   2593 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 2652

Qy    267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
          |||||||||:::||||||:::   |||||||||||||||||||||||||:::||||||
Db   2653 AAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 2712

Qy    287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
          ||||||:::||||||||||||:::|||:::      ||||||   ||||||   :::   |||
Db   2713 CGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTG 2772

Qy    307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
          |||         :::||||||:::   :::||||||   |||   |||      ::::::|||
Db   2773 CAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTC 2832

Qy    327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
          :::   |||   :::||||||:::||||||      ||||||||||||||||||||:::
Db   2833 ACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTG 2892

Qy    347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
                   |||:::   |||   :::   :::||||||:::|||:::   :::|||
Db   2893 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGT 2952

Qy    367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
          |||||||||:::|||:::               ||||||:::|||:::||||||||
Db   2953 GCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCA 3012

Qy    387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
          |||||||||:::         ::::::||||||   |||   ||||||||   |||   |||
Db   3013 GCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAA 3072

Qy    407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
          |||         |||:::|||   |||||||||:::   |||   ||||||   |||   |||
Db   3073 GGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTAT 3132

Qy    427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
          |||   |||||||||:::   :::|||:::||||||||||||   :::::::::|||:::
Db   3133 CCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAG 3192

Qy    447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
          |||:::|||   ||||||||||||||||||||||||||||   ||||||||||||:::
Db   3193 ACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAG 3252

Qy    467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
          |||   |||||||||:::   |||:::||||||   ||||||   :::   |||   |||
Db   3253 CGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTG 3312

Qy    487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
          ||||||||||||||||||:::   :::   ||||||   ||||||||:::||||||:::
Db   3313 AATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGAC 3372

Qy    507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
          ||||||||||||||||||||||||||||||||||||||||||         :::|||
Db   3373 TGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAG 3432
```

```
Qy     527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
           |||      ||||||      |||||||||||:::|||||||||   ||||||   |||
Db    3433 ATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAA 3492

Qy     547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
           |||:::|||:::||||||   ||||||   ||||||||||||||||||||||||||:::
Db    3493 TATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATT 3552

Qy     567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
           ||||||||||||||||||:::::::::|||   ||||||||||||||||||||||||||||
Db    3553 GCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCA 3612

Qy     587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
           |||||||||   :::||||||||||||||||||| . |||||||||||||||||   |||
Db    3613 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 3672

Qy     607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
           |||||||||:::|||:::   ||||||||||||   ||||||||||||||||||||||||
Db    3673 CGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTG 3732

Qy     627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
           |||      ||||||::::::||||||   ||||||||||||:::||||||      :::
Db    3733 GTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAA 3792

Qy     647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
              |||||||||   :::|||:::   |||:::
Db    3793 GGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3825
```

Search completed: March 30, 2003, 03:05:40
Job time : 735.062 secs

GenCore version 5.1.4_p5_4578


```
OM protein - nucleic search, using frame_plus_p2n model

Run on:          March 30, 2003, 03:22:44 ; Search time 4532.72 Seconds
                                            (without alignments)
                                            2354.618 Million cell updates/sec

Title:           US-09-873-409-1
Perfect score:   3323
Sequence:        1 MLAEKGAHAELMAKRGLYYS..........QELLRNRDIYFKLVNAQSVQ 659

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters:      32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115420_19240/app_query.fasta_1.7544
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_30544_@runat_27032003_115420_19240 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       EST:*
                 1:  em_estba:*
                 2:  em_esthum:*
                 3:  em_estin:*
                 4:  em_estmu:*
                 5:  em_estov:*
                 6:  em_estpl:*
                 7:  em_estro:*
                 8:  em_htc:*
                 9:  gb_est1:*
                 10:  gb_est2:*
                 11:  gb_htc:*
                 12:  gb_est3:*
                 13:  gb_est4:*
                 14:  gb_est5:*
                 15:  em_estfun:*
                 16:  em_estom:*
                 17:  gb_gss:*
                 18:  em_gss_hum:*
                 19:  em_gss_inv:*
                 20:  em_gss_pln:*
                 21:  em_gss_vrt:*
                 22:  em_gss_fun:*
                 23:  em_gss_mam:*
                 24:  em_gss_mus:*
                 25:  em_gss_other:*
                 26:  em_gss_pro:*
                 27:  em_gss_rod:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1008	30.3	943	9	AL520322	AL520322 AL520322
	2	834	25.1	998	14	BM904842	BM904842 AGENCOURT
	3	780	23.5	2676	11	AK014319	AK014319 Mus muscu
	4	770.5	23.2	1019	12	BG248052	BG248052 602359987
	5	753	22.7	726	12	BG293345	BG293345 602390738
	6	751	22.6	944	12	BF796582	BF796582 602258463
	7	743.5	22.4	886	17	BH139685	BH139685 ENTNA47TR
	8	740.5	22.3	871	17	AZ682350	AZ682350 ENTKB16TF
c	9	734.5	22.1	932	17	AZ670821	AZ670821 ENTJN69TF
	10	728.5	21.9	947	17	AZ683753	AZ683753 ENTIL96TF
	11	726.5	21.9	886	17	AZ540627	AZ540627 ENTEQ18TF
c	12	723.5	21.8	880	17	AZ687805	AZ687805 ENTLU52TF
	13	721.5	21.7	939	14	BQ720763	BQ720763 AGENCOURT
	14	720.5	21.7	897	17	AZ541090	AZ541090 ENTDS67TR
	15	711.5	21.4	913	17	BH155700	BH155700 ENTRO54TR
	16	709.5	21.4	834	17	AZ548312	AZ548312 ENTFG07TR
c	17	708.5	21.3	853	17	AZ679807	AZ679807 ENTHI67TR
	18	705.5	21.2	823	17	AZ532602	AZ532602 ENTCR50TR
	19	697.5	21.0	861	14	BQ717101	BQ717101 AGENCOURT
	20	691	20.8	780	12	BG587938	BG587938 EST489713
	21	690.5	20.8	899	17	BH154857	BH154857 ENTRO47TF
c	22	689.5	20.7	891	17	AZ682250	AZ682250 ENTKT68TR
	23	688	20.7	1123	11	AY108485	AY108485 Zea mays
	24	683.5	20.6	1341	11	AY108285	AY108285 Zea mays
	25	683	20.6	600	13	BM486593	BM486593 pgm2n.pk0
	26	682.5	20.5	919	17	AZ690701	AZ690701 ENTIY68TR
	27	678	20.4	926	14	BQ123477	BQ123477 EST609053
	28	677.5	20.4	795	12	BF313560	BF313560 601900192
	29	669	20.1	547	13	BJ030711	BJ030711 BJ030711
c	30	667	20.1	937	17	BH150760	BH150760 ENTQG42TR
c	31	663	20.0	835	13	BJ349604	BJ349604 BJ349604
	32	662	19.9	636	10	AV962688	AV962688 AV962688
c	33	661	19.9	747	13	BJ445774	BJ445774 BJ445774
	34	659.5	19.8	628	12	BG080311	BG080311 H3052B06-
	35	659	19.8	852	17	BH720383	BH720383 BOHZZ94TF
	36	657	19.8	715	14	BQ869512	BQ869512 QGD6G09.y
	37	657	19.8	785	10	AV709991	AV709991 AV709991
	38	655.5	19.7	787	12	BG584063	BG584063 EST485823
	39	653	19.7	609	10	AV986144	AV986144 AV986144
c	40	651.5	19.6	921	17	AZ687628	AZ687628 ENTIR96TR
c	41	651	19.6	746	13	BJ380617	BJ380617 BJ380617
c	42	650	19.6	729	13	BJ353293	BJ353293 BJ353293
	43	647	19.5	1813	11	U66688	U66688 Homo sapien
c	44	641	19.3	820	13	BJ356540	BJ356540 BJ356540
c	45	639.5	19.2	842	17	AZ671925	AZ671925 ENTMN93TR

ALIGNMENTS

```
RESULT 1
AL520322
LOCUS       AL520322                 943 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL520322 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB006YC15 5
            prime, mRNA sequence.
ACCESSION   AL520322
VERSION     AL520322.1  GI:12783815
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 943)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
```

```
  JOURNAL     Unpublished (2001)
COMMENT       Contact: Genoscope
              Genoscope - Centre National de Sequencage
              BP 191 91006 EVRY cedex - France
              Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES               Location/Qualifiers
     source            1..943
                       /organism="Homo sapiens"
                       /db_xref="taxon:9606"
                       /clone="CS0DB006YC15"
                       /clone_lib="LTI_NFL004_NBC2"
                       /sex="male"
                       /tissue_type="neuroblastoma cells"
                       /lab_host="DH10B"
                       /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                       was primed with a NotI-oligo(dT) primer. Five prime end
                       enriched, double-stranded cDNA was digested with Not I and
                       cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                       vector.  Library was normalized. Library was constructed
                       by Life Technologies.  Contact : Feng Liang Life
                       Technologies, a division of Invitrogen 9800 Medical Center
                       Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
                       8371 Email : fliang@lifetech.com URL :
                       http://fulllength.invitrogen.com"
BASE COUNT      253 a    225 c    253 g    211 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              3.17e-110      Length:        943
Score:                  1008.00        Matches:       194
Percent Similarity:     80.40%         Conservative:  48
Best Local Similarity:  64.45%         Mismatches:    58
Query Match:            30.33%         Indels:        1
DB:                     9              Gaps:          0

US-09-873-409-1 (1-659) x AL520322 (1-943)

Qy    358 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 377
          ::::::|||:::|||:::   :::|||||||||||:::|||:::             |||
Db      5 TTATTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCT 64

Qy    378 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 397
          |||:::|||:::|||||||||      |||||||||:::      ::::::|||||
Db     65 CCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACC 124

Qy    398 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 417
          |||   |||||||||   |||   ||||||      |||:::|||   |||||||||:::
Db    125 CCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTC 184

Qy    418 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 437
             |||   ||||||   |||   ||||||   |||||||||:::   :::|||:::|||
Db    185 ACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGA 244

Qy    438 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 457
          |||||||||   :::::::::|||:::|||:::|||   |||||||||||||||||||||
Db    245 CTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGG 304

Qy    458 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 477
          |||||||||   ||||||||||||:::|||   ||||||||||:::   |||:::||||||
Db    305 AAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTG 364

Qy    478 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 497
             ||||||   :::   |||   ||||||||||||||||||||||:::   :::   |||
Db    365 CTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATC 424

Qy    498 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 517
          |||   |||||||||:::||||||:::||||||||||||||||||||||||||||||||||
Db    425 GTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGAC 484

Qy    518 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 537
          |||||||||||||||      :::||||||      ||||||      ||||||||||||
Db    485 AACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACAT 544

Qy    538 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 557
          :::|||||||||   ||||||   ||||||:::|||:::||||||   ||||||   |||
Db    545 GCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAG 604

Qy    558 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 577
          |||||||||||||||||||||||||||:::|||||||||||||||||||:::::::::|||
Db    605 CTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCAT 664

Qy    578 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 597
          |||||||||||||||||||||||||||||||||      :::||||||||||||||||||
Db    665 ATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAA 724

Qy    598 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 617
             ||||||||||||||||||   ||||||||||||:::|||:::   ||||||||||||
Db    725 GAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCC 784

Qy    618 AlaIleGlnAsnAlaAspLeuIleVal-ValLeuHisAsnGlyLysIleLysGluGlnGl 637
             |||||||||||||||||||||||| |||      ||||||::::::||||||   ||
Db    785 ACCATCCAGAATGCAGACTTAATAGTGTGTGTTTCAGAATGGCAGAGTCAAGGAGCATGG 844

Qy    637 yThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSe 657
           ||||||||:::||||||      :::   |||||||||   :::|||:::   |||::
Db    845 CACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGC 904

Qy    657 r 657
          :
Db    905 T 905

RESULT 2
BM904842
LOCUS       BM904842                998 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6699581 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5557655
            5', mRNA sequence.
ACCESSION   BM904842
VERSION     BM904842.1  GI:19355221
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 998)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12279  row: f  column: 24
            High quality sequence stop: 738.
FEATURES             Location/Qualifiers
     source          1..998
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5557655"
                     /clone_lib="NIH_MGC_72"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      271 a    223 c    248 g    256 t
ORIGIN

Alignment Scores:
Pred. No.:              3.26e-89       Length:        998
Score:                  834.00         Matches:       164
```

```
Percent Similarity:     74.12%          Conservative:   68
Best Local Similarity:  52.40%          Mismatches:     77
Query Match:            25.10%          Indels:         4
DB:                     14              Gaps:           0

US-09-873-409-1 (1-659) x BM904842 (1-998)

Qy    182 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 201
          |||::::::|||||||||:::   :::|||:::||||||   |||||||||   ||||||
Db     61 GATGTGAGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCC 120

Qy    202 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 221
             |||   ||||||::::::||||||   |||||||||:::   |||:::|||||||||
Db    121 AATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAAT 180

Qy    222 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 241
                |||:::|||      :::||||||||||||||||||||||||::::::|||
Db    181 ATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTG 240

Qy    242 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 261
          |||:::|||:::|||   |||::::::|||:::   |||::::::|||         :::
Db    241 TTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTG 300

Qy    262 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 281
          :::|||   |||   |||||||||:::||||||:::   |||||||||||||||||||||
Db    301 TCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAA 360

Qy    282 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 301
          |||:::||||||   ||||||:::||||||||||||:::|||:::   ||||||   |||
Db    361 GCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATG 420

Qy    302 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 321
          |||   :::   ||||||      :::||||||:::   :::||||||   |||   |||
Db    421 TATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGA 480

Qy    322 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 341
                ::::::|||:::   |||   :::||||||:::||||||      |||||||||
Db    481 ATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTT 540

Qy    342 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 361
          ||||||||||||:::            |||:::   |||   :::   :::||||||:::
Db    541 GGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCA 600

Qy    362 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 381
          |||:::   :::||||||||||||:::|||:::            ||||||:::|||:::
Db    601 GCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCC 660

Qy    382 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 401
          |||||||||      |||||||||:::      ::::::||||||   |||   ||||||
Db    661 AAAGCCAAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGAC 720

Qy    402 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 421
          |||   |||   ||||||      |||:::|||   |||||||||:::   |||
Db    721 AGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGGTGA 780

Qy    422 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer-LeuSe 441
          |||   |||   ||||||   |||||||||:::   :::|||:::||||||||| |||
Db    781 GTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCCTGGA 840

Qy    441 rIleGluArgGlyLysThrValAlaPheVal-GlySerSerGlyCysGlyLysSerThrS 461
           :::::::::|||:::|||:::|||   ||| |||||||||||||||||||||||||||
Db    841 GGTGAAGAAGGGGCAGACCCTGGCTCTGGTGGGGCAGCAGTGGCTGTGGGAAGAGCACAG 900

Qy    461 er-ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480
             ||||||||||||:::|||:::|||:::|||:::      ::::::|||   ||||||
Db    901 TGGGTCCAGCTCCTGGAGCGGGTCTACAACCCCTTGGCAAGGAAAATGCTGCTTGATGGC 960

Qy    481 ValAspAlaLys-GluLeuAsnValGlnTrp 490
                   |||    |||:::|||   |||
Db    961 AAAAGAATAAAACCGACTGGATGTTCCATGG 991

RESULT 3
AK014319
```

```
LOCUS       AK014319             2676 bp    mRNA    linear   HTC 19-JAN-2002
DEFINITION  Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length
            enriched library, clone:3222401P09:ATP-binding cassette, sub-family
            B (MDR/TAP), member 8, full insert sequence.
ACCESSION   AK014319
VERSION     AK014319.1  GI:12852089
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (strain:C57BL/6J) 14, 17 days embryo head cDNA to
            mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:3222401P09.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409 (6821), 685-690 (2001)
  MEDLINE   21085660
   PUBMED   11217851
REFERENCE   5  (bases 1 to 2676)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C.,
            Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,
            Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F.,
            Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,
            Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,
            Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
            Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,
```

```
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,
            Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and
            Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Gemomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. Second
            strand cDNA was prepared with the primer adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cleaved
            with BamHI and XhoI. cDNA of size comprised between 3 and 7 kb was
            selected before cloning. Vector: a modified pBluescript KS(+) after
            bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'
            end: BamHI. Host: DH10B.
FEATURES             Location/Qualifiers
     source          1..2676
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:3222401P09"
                     /db_xref="MGD:MGI:1907488"
                     /db_xref="taxon:10090"
                     /clone="3222401P09"
                     /tissue_type="head"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="14, 17 days embryo"
     gene            1..2676
                     /gene="Abcb8"
     CDS             72..2225
                     /gene="Abcb8"
                     /note="ATP-binding cassette, sub-family B (MDR/TAP),
                     member 8
                     data source:MGD, source key:MGI:1351667, evidence:ISS
                     putative"
                     /codon_start=1
                     /protein_id="BAB29270.1"
                     /db_xref="GI:12852090"
                     /db_xref="MGD:MGI:1351667"
                     /translation="MLVHLFRFGIRGGPVPGWSLQSLRFQTFSAARSSDDRLSSHLLR
                     TVAQLRVQLRAHLPRAPPASHWSPSAWCWVGGTLVVPAVLWQHPRLCLIALCEAKESP
                     PAQPTRAPELRFNWKLFWHFLHPHLLALGAAIVLALGAALVNVQIPLLLGQLVEIVAK
                     YTRDHMGSFVSESRKLSVQLLLLYGVQGLLTFGYLVLLSHIGERMAMDMRKALFSSLL
                     RQDIAFFDAKKTGQLVSRLTTDVQEFKSSFKLVISQGLRSCTQVIGSLVSLSMLSPRL
                     TLMLAVVTPALMGVGTLMGSGLRKLSRQCQEQIARATGVADEALGNVRTVRAFAMEKR
                     EEERYQAELESCCCKAEELGRGIALFQGLSNIAFNCMVLGTLFIGGSLVAGQQLKGGD
                     LMSFLVASQTVQRSMASLSVLFGQVVRGLSAGARVFEYMALSPVIPLTGGYCIPNKDI
                     RGSITFQNVTFSYPCRPGFNVLKDFTLKLPSGKIVALVGQSGGGKTTVASLLERFYDP
                     EAGSVTLDGHDLRTLNPSWLRGQVIGFISQEPVLFATTIMENIRFGKLDASDEEVYTA
                     AREANAHEFISSFPDGYSTVVGERGTTLSGGQKQRLAIARALIKQPTVLILDEATSAL
                     DAESERVVQEALDRASAGRTVLVIAHRLSTVRAAHSIIVMANGQVCEAGTHEELLKKG
                     GLYSELIRRQTLDASLTSTPPAEKPEDPKSCQSKA"
     polyA_signal    2653..2658
                     /gene="Abcb8"
                     /note="putative"
     polyA_site      2676
                     /gene="Abcb8"
                     /note="putative"
BASE COUNT      544 a    785 c    756 g    591 t
```

```
ORIGIN

Alignment Scores:
Pred. No.:              5.46e-82      Length:        2676
Score:                  780.00        Matches:       207
Percent Similarity:     51.28%        Conservative:  113
Best Local Similarity:  33.17%        Mismatches:    250
Query Match:            23.47%        Indels:        54
DB:                     11            Gaps:          14

US-09-873-409-1 (1-659) x AK014319 (1-2676)

Qy       63 LysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLys 82
              :::|||:::|||||     ::::::            ||||||
Db      360 GAGGCCAAAGAGTCCCCTCCTGCCCAGCCCACACGTGCCCCAGAG--------------- 404

Qy       83 IleLeuLysLeuAsn---LysProGluTrpProPhe------------------------ 93
              |||:::   |||   |||     |||   |||
Db      405 ---CTGCGCTTTAATTGGAAGCTCTTCTGGCATTTCCTTCACCCCCACCTGCTGGCCTTG 461

Qy       94 ---------ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProVal 110
                     :::||||||         |||:::::::::|||      :::   |||:::
Db      462 GGGGCAGCCATTGTGCTGGCCTTAGGTGCGGCACTAGTGAATGTGCAGATC---CCCTTG 518

Qy      111 PheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeu 130
                                 :::|||:::       :::       :::         :::
Db      519 CTCCTGGGCCAGCTGGTGGAGATTGTCGCCAAGTACACGAGGGACCACATGGGGAGTTTC 578

Qy      131 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 150
                  ::::::         |||:::      ::::::|||
Db      579 GTGTCTGAGTCCCGTAAGCTCAGCGTCCAGCTGCTCCTACTG------------------ 620

Qy      151 SerTyrPheMetGlnGlyLeu--------------------PheTyrGlyArgAlaGly 163
               |||   :::||||||||                                     |||
Db      621 ---TACGGTGTTCAGGGACTGCTGACCTTTGGATACCTAGTGCTGCTGTCCCACATTGGT 677

Qy      164 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 183
            |||   :::   |||   :::|||      |||   :::::|||   ||||||||
Db      678 GAGCGCATGGCCATGGACATGCGGAAAGCCCTTTTCAGCTCCCTGCTCCGGCAAGACATT 737

Qy      184 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 203
             |||:::||||||   |||:::      ||||||   |||   :::   |||      |||
Db      738 GCTTTCTTTGATGCCAAAAAG------ACAGGGCAGCTAGTGAGTCGCTTGACTACTGAT 791

Qy      204 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 223
            :::   :::   :::   :::                  |||            ::::::   |||
Db      792 GTGCAAGAATTCAAGTCATCCTTCAAGCTTGTCATCTCCCAGGGACTGCGCAGCTGCACC 851

Qy      224 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 243
               :::      |||::::::         :::            :::|||   ::::::
Db      852 CAGGTGATTGGTAGCCTGGTGTCCCTGTCTATGCTGTCCCCTCGCCTTACCCTGATGCTG 911

Qy      244 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 263
               :::   |||   |||         |||   :::      :::   :::
Db      912 GCTGTCGTCACACCCGCCCTCATGGGAGTGGGCACCCTGATGGGCTCAGGCCTCCGAAAG 971

Qy      264 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 283
               :::   :::   ::::::::::::         |||      :::|||   ||||||||
Db      972 CTGTCTCGCCAGTGTCAGGAGCAGATTGCCAGGGCAACAGGTGTAGCAGATGAGGCCCTT 1031

Qy      284 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 303
               |||:::||||||:::   :::            ||||||   |||:::   |||:::
Db     1032 GGCAATGTTCGGACTGTGCGGGCCTTCGCGATGGAGAAGAGGGAGGAGGAACGCTATCAA 1091

Qy      304 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 323
                 |||:::            :::         ||||||:::   :::|||
Db     1092 GCAGAACTGGAG-----------TCATGCTGCTGTAAAGCAGAAGAACTGGGCAGGGGC 1139

Qy      324 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyr---------AlaAlaGlyPheArg 340
               |||                     |||:::                   |||
Db     1140 ATCGCCTTGTTCCAAGGGCTCTCCAACATCGCTTTCAACTGTATGGTCTTGGGCACCCTG 1199

Qy      341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGly---------MetPhe 357
```

```
                |||              :::   ||||||:::              |||            |||
Db    1200 TTCATTGGGGGCTCCCTTGTGGCTGGACAGCAGCTGAAAGGGGGAGACCTCATGTCCTTC 1259

Qy     358 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 377
           :::|||   :::               :::||||||                |||||| 
Db    1260 CTGGTGGCTTCCCAGACAGTACAGAGGTCTATGGCC------AGCCTCTCTGTCCTGTTT 1313

Qy     378 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 397
              :::   :::      |||       |||   :::|||      :::
Db    1314 GGTCAGGTGGTACGTGGGCTGAGTGCGGGAGCCCGAGTCTTCGAATACATGGCCCTGAGC 1373

Qy     398 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 417
           |||   |||                                          |||::::::
Db    1374 CCTGTCATCCCATTGACCGGGGGCTACTGCATCCCCAACAAGGACATTCGTGGTTCCATC 1433

Qy     418 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 437
              |||:::   |||:::|||   ||||||||||||||           :::|||:::
Db    1434 ACCTTCCAAAATGTCACCTTCAGCTACCCCTGCAGACCTGGCTTCAATGTGCTCAAGGAC 1493

Qy     438 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 457
              :::|||   :::      ||||||   ||||||   ||||||   ||||||   |||
Db    1494 TTCACCCTGAAGCTGCCCTCTGGCAAGATTGTGGCTCTTGTGGGCCAGTCTGGGGGAGGA 1553

Qy     458 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 477
           |||:::|||         |||||||:::|||   |||||||||      |||   |||
Db    1554 AAGACCACAGTTGCCTCCTTGCTGGAACGCTTCTATGACCCTGAAGCTGGCTCGGTGACG 1613

Qy     478 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln---IleAla 496
              ||||||   |||   :::   ||||||      |||||||||   |||   |||
Db    1614 TTGGATGGGCATGACCTGCGAACTCTCAACCCCTCCTGGCTCCGGGGCCAGGTCATAGGT 1673

Qy     497 IleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGly 516
              :::   ||||||||||||||||||         :::|||   |||||||||   :::|||
Db    1674 TTCATCAGCCAGGAGCCAGTCCTGTTTGCAACAACCATCATGGAGAATATCCGATTTGGG 1733

Qy     517 ------AspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 534
                    |||   |||               :::|||:::      ||||||      |||
Db    1734 AAGCTGGATGCTTCCGAT------------GAAGAGGTGTACACAGCTGCACGAGAAGCC 1781

Qy     535 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 554
           |||   |||   ||||||            |||:::   |||:::|||   ||||||   :::
Db    1782 AATGCCCACGAGTTCATCAGCAGCTTCCCCGATGGCTACAGCACTGTGGTTGGTGAGCGG 1841

Qy     555 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 574
           |||      |||||||||||||||||||||||||||||||||||||||||||||:::::: 
Db    1842 GGCACAACCTTGTCTGGTGGCCAGAAGCAGCGCCTAGCCATCGCACGTGCCCTCATCAAG 1901

Qy     575 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 594
           :::|||   :::|||:::|||||||||||||||||||||||||||   :::||||||:::
Db    1902 CAGCCCACAGTGCTGATCCTGGACGAGGCCACCAGTGCGCTAGATGCAGAATCCGAGAGG 1961

Qy     595 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 614
           |||||||||   |||||||||:::|||         |||||||||   ||||||:::   |||
Db    1962 GTGGTACAGGAGGCCCTGGACCGGGCCAGTGCTGGCCGCACCGTGTTGGTCATTGCCCAC 2021

Qy     615 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 634
           |||||||||   ::::::   |||      |||:::|||:::      ||||||::::::
Db    2022 CGGCTTAGTACTGTCCGTGCAGCCCACTCCATCATTGTCATGGCCAATGGCCAAGTCTGT 2081

Qy     635 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 654
           |||   |||||||||:::|||||||||:::               :::|||   :::|||:::
Db    2082 GAGGCTGGGACCCACGAAGAACTCCTTAAAAAAGGCGGGCTCTATTCAGAGCTTATCCGG 2141

Qy     655 AlaGlnSerVal 658
              |||::::::
Db    2142 AGACAAACCCTG 2153
```

RESULT 4
BG248052

```
LOCUS       BG248052            1019 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  602359987F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4488404 5',
            mRNA sequence.
ACCESSION   BG248052
VERSION     BG248052.1  GI:12757867
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1019)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10334  row: n  column: 21
            High quality sequence stop: 650.
FEATURES             Location/Qualifiers
     source          1..1019
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4488404"
                     /clone_lib="NCI_CGAP_Mam1"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="10 months, virgin"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
BASE COUNT      276 a    262 c    283 g    196 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:               1.52e-81       Length:          1019
Score:                   770.50         Matches:         154
Percent Similarity:      79.17%         Conservative:    36
Best Local Similarity:   64.17%         Mismatches:      47
Query Match:             23.19%         Indels:          3
DB:                      12             Gaps:            1

US-09-873-409-1 (1-659) x BG248052 (1-1019)

Qy     414 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 433
           |||||||||::::::|||         |||   |||   ||||||   ||||||      :::
Db       4 GAAGGAAATGTGCAATTTAGTGGAGTCGTGTTCAACTATCCCACCCGACCCAGCATCCCA 63

Qy     434 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 453
           :::|||:::||||||||||||      :::::::::|||:::|||:::|||   |||||||||
Db      64 GTGCTTCAGGGGCTGAGCCTTGAGGTGAAGAAGGGCCAGACGCTGGCCCTGGTGGGCAGC 123

Qy     454 SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGln 473
           ||||||||||||||||||||||||   ||||||||||||:::|||   |||||||||:::
Db     124 AGTGGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAGCGCTTCTACGACCCCATGGCT 183

Qy     474 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 493
           |||   |||      ||||||   :::      |||:::|||||||||||||||||||||:::
Db     184 GGATCAGTGTTTCTAGATGGCAAAGAAATAAAGCAACTGAATGTCCAGTGGCTCCGAGCA 243

Qy     494 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 513
           |||:::   ||||||   |||||||||:::||||||:::||||||||||||||||||||||
Db     244 CAGCTGGGCATTGTGTCCCAAGAGCCCATTCTCTTTGACTGCAGCATCGCAGAGAACATT 303

Qy     514 AlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla 533
           |||||||||||||||||||||||||||         :::||||||      ||||||
Db     304 GCCTACGGAGACAACAGCCGGGTCGTGTCTTATGAGGAGATTGTGAGGGCAGCCAAGGAG 363

Qy     534 AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeu 553
```

```
Db    364 GCCAACATCCACCAGTTCATCGACTCGCTACCTGATAAATACAACACCAGAGTAGGAGAC 423

Qy    554 LysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeu 573
Db    424 AAAGGCACTCAGCTGTCGGGTGGGCAGAAGCAGCGCATCGCCATCGCACGCGCCCTCGTC 483

Qy    574 GlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGlu 593
Db    484 AGACAGCCTCACATTTTACTTCTGGACGAAGCAACATCAGCTCTGGATACAGAAAGTGAA 543

Qy    594 LysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThr 613
Db    544 AAGGTTGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACCTGCATTGTGATCGCT 603

Qy    614 HisArgLeuSerAlaIleGlnAsnAlaAspLeuIle-ValValLeuHis-AsnGlyLysI 633
Db    604 CACCGCTGTCCACCATCCCAGAACGCGGACTTGATCGGTGGTGATTCAAGAACCGGAAGG 663

Qy    633 leLys---GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 650
Db    664 TCAAGGGAGCACGGGCACCCCAACAGCAAGTGGTGGCGCAGAAGGGCATCTACTTC 719
```

```
RESULT 5
BG293345
LOCUS       BG293345                726 bp    mRNA    linear   EST 21-FEB-2001
DEFINITION  602390738F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4502552 5',
            mRNA sequence.
ACCESSION   BG293345
VERSION     BG293345.1  GI:13052943
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 726)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10371  row: l  column: 09
            High quality sequence stop: 658.
FEATURES             Location/Qualifiers
     source          1. .726
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4502552"
                     /clone_lib="NIH_MGC_94"
                     /tissue_type="retina"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 3.3 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      189 a    220 c    190 g    127 t
ORIGIN

Alignment Scores:
Pred. No.:               1.11e-79      Length:         726
Score:                   753.00        Matches:        152
Percent Similarity:      84.29%        Conservative:   25
Best Local Similarity:   72.38%        Mismatches:     30
Query Match:             22.66%        Indels:         3
DB:                      12            Gaps:           0
```

```
US-09-873-409-1 (1-659) x BG293345 (1-726)

Qy    448 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 467
Db      2 CTGGCCCTGGTGGGCAGCAGTGGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAGCGC 61

Qy    468 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 487
Db     62 TTCTATGACCCCATGGCTGGATCAGTGCTCTTAGATGGTCAAGAAGCAAAGAAACTCAAT 121

Qy    488 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 507
Db    122 GTCCAGTGGCTCCGAGCTCAACTGGGCATTGTGTCCCAGGAACCCATTCTCTTTGACTGC 181

Qy    508 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 527
Db    182 AGCATCGCAGAGAACATCGCCTATGGAGACAACAGCCGGGTCGTGCCTCATGATGAGATT 241

Qy    528 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 547
Db    242 GTGAGGGCAGCCAAGGAGGCCAACATCCACCCCTTCATCGAGACGCTGCCCCAAAAATAT 301

Qy    548 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 567
Db    302 AACACAAGAGTAGGAGACAAGGGGACGCAGCTCTCTGGGGGCCAGAAGCAGAGGATTGCC 361

Qy    568 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 587
Db    362 ATCGCCCGAGCCCTCATCAGACAGCCTCGGGTCCTACTGCTGGATGAAGCCACGTCAGCT 421

Qy    588 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 607
Db    422 CTGGATACTGAGAGTGAAAAGGTTGTCCAGGAAGCACTGGACAAAGCCAGGGAAGGCCGC 481

Qy    608 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla-AspLeuIleValVa 627
Db    482 ACCTGCATTGCGATCGCTCACCGCCTGTCCACCATCCAGAACGCGGCACTTGATCGTGGT 541

Qy    627 lLeuHisAsnGlyLysIleLysGluGlnGlyThr-HisGlnGluLeuLeuArgAsnArgA 647
Db    542 GATTGAGAACGGCAAGGTCAAGGAGCACGGCACCCCACCAGCAGCTGCTGGCGCAGAAGG 601

Qy    647 spIleTyrPhe-LysLeuValAsn 654
Db    602 GCATCTATTTCTCAACTGGTCAAC 625
```

```
RESULT 6
BF796582
LOCUS       BF796582                944 bp    mRNA    linear   EST 12-JAN-2001
DEFINITION  602258463F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4341710 5',
            mRNA sequence.
ACCESSION   BF796582
VERSION     BF796582.1  GI:12101636
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 944)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis Staudt, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9955  row: n  column: 15
```

```
           High quality sequence start: 9
           High quality sequence stop: 669.
FEATURES             Location/Qualifiers
     source          1..944
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4341710"
                     /clone_lib="NIH_MGC_85"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.867 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      265 a    241 c    237 g    201 t
ORIGIN

Alignment Scores:
Pred. No.:              3e-79          Length:         944
Score:                  751.00         Matches:        166
Percent Similarity:     74.62%         Conservative:   31
Best Local Similarity:  62.88%         Mismatches:     60
Query Match:            22.60%         Indels:         9
DB:                     12             Gaps:           2

US-09-873-409-1 (1-659) x BF796582 (1-944)

Qy    380 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 399
          |||:::|||||||||      ||||||||||||||    |||    |||:::::::|||
Db     39 TATGCTAAAGCTAAGCTGTCTGCAGCCCACTTATTCATGCTGTTTGAAAGACAACCTCTG 98

Qy    400 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 419
          |||||||||   |||:::||||||   |||||||||          ||||||||:::    |||
Db     99 ATTGACAGCTACAGTGAAGAGGGGCTGAAGCCTGATAAATTTGAAGGAAATATAACATTT 158

Qy    420 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 439
                ||||||    |||   ||||||   |||     :::|||   :::|||:::|||||||||
Db    159 AATGAAGTCGTGTTCAACTATCCCACCCGAGCAAACGTGCCAGTGCTTCAGGGGCTGAGC 218

Qy    440 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 459
          |||   :::::::::|||:::|||:::|||   ||||||||||||||||||||||||||||
Db    219 CTGGAGGTGAAGAAAGGCCAGACACTAGCCCTGGTGGGCAGCAGTGGCTGTGGGAAGAGC 278

Qy    460 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 479
          |||   ||||||||||||:::|||   ||||||||||:::   |||   ||||||    |||
Db    279 ACGGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCGGGGACAGTGCTTCTCGAT 338

Qy    480 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 499
          |||   :::||||||:::||||||||||||||||||||||:::|||:::   ||||||
Db    339 GGTCAAGAAGCAAAGAAACTCAATGTCCAGTGGCTCAGAGCTCAACTCGGAATCGTGTCT 398

Qy    500 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 519
          |||||||||:::||||||:::||||||||||||||||||||||||||||||||||||||||
Db    399 CAGGAGCCTATCCTATTTGACTGCAGCATTGCCGAGAATATTGCCTATGGAGACAACAGC 458

Qy    520 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 539
          |||||||||      |||||||||         ||||||   ||||||||||||||||    |||
Db    459 CGGGTTGTATCACAGGATGAAATCGTGAGTGCAGCCAAAGCTGCCAACATACATCCTTTC 518

Qy    540 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 559
          ||||||    ||||||    ||||||    |||:::||||||    ||||||   |||||||||
Db    519 ATCGAGACGTTACCCCACAAATATGAAACAAGAGTGGGAGATAAGGGGACTCAGCTCTCA 578

Qy    560 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLys-ProLysIleLe 579
          ||||||    ||||||||:::||||||||||||||||||:::::::::  |||:::|||||
Db    579 GGAGGTCACAAACAGAGGATTGCTATTGCCCGAGCCCTCATCAGACAACCCTCAAATCCT 638

Qy    579 uLeuLeuAspGluAlaThrSerAla-LeuAspAsnAspSerGluLysValVal------G 597
          ||||||||||||||||||||||||  ||||||                 |||        :
Db    639 CCTGTTGGATGAAGCTACATCAGCCTCTGGATACCTGACAAGTGAACACGGTTGTCCCAA 698

Qy    597 lnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuS 617
                                     ::                 |||              ||||||||||||||:::     ||||||
Db    699 GAAGCCCCTGTCACACAGCAGAAGAGCGCCGCACCTGCCTTGTGATTGCTCACCGCTGTC 758

Qy    617 erAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnG 637
               ||||||||||||||   |||||||||           |||::::::|||
Db    759 AAC--ATCCAGAATGCAGACTT-ATAGTGGTGTTCCCAACGGGAGAAGTCAAG------- 808

Qy    637 lyThrHis 639
              ||||||
Db    809 --ACACAT 814

RESULT 7
BH139685
LOCUS       BH139685                 886 bp    DNA     linear   GSS 07-AUG-2001
DEFINITION  ENTNA47TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   BH139685
VERSION     BH139685.1  GI:15098746
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
  ORGANISM  Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 886)
  AUTHORS   Loftus,B., Wang,Z., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library (2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Reverse
            Class: shotgun
            High quality sequence start: 18
            High quality sequence stop: 851.
FEATURES             Location/Qualifiers
     source          1..886
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a  method described by Clark and Diamond (Clark,
                     C.G., and Diamond, L.S.  (1993) Entamoeba histolytica: a
                     method for isolate identification.  Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + i method used for
                     the library  construction is described in detail in Smith,
                     H.O. and Venter, J.C.  (Making small insert libraries for
                     whole genome shotgun sequencing  projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin  and B.
                     Barell, Oxford University Press, 1999)."
BASE COUNT      348 a    105 c    176 g    257 t
ORIGIN

Alignment Scores:
Pred. No.:              2.16e-78       Length:         886
Score:                  743.50         Matches:        151
Percent Similarity:     73.43%         Conservative:   48
Best Local Similarity:  55.72%         Mismatches:     71
Query Match:            22.37%         Indels:         2
DB:                     17             Gaps:           1

US-09-873-409-1 (1-659) x BH139685 (1-886)

Qy    387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
```

```
Db     20 GCAAAGAATACTTATGATGTTATTGATAGAAAACCACT-ATTGATTGTTATAGTGAAGAA 78

Qy    407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
Db     79 GGTGAAACATTTAATGATGTTAAAGGTGAAATTGAATTTAAAGACATTTGTTTCAGATAT 138

Qy    427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db    139 CCAACAAGACCAGACAATTCTGTCTTGAAAGGTATTTCATTCAAAGTAGAACAAGGAAAA 198

Qy    447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
Db    199 ACTGTTGCATTAGTAGGAGCATCAGGATGTGGTAAATCAACATCAGTTCAGTTGATTGAA 258

Qy    467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db    259 AGATTTTATGATCCAACACATGGAGATGTATTATTAGATGGACATAATATCAAAGATTTG 318

Qy    487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db    319 AATATTCATTTCTTAAGAAGTCAAATTGGAATGGTAGGACAAGAACCAGTATTATTTGCT 378

Qy    507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal---ValProLeuAsp 525
Db    379 GAAAGTGTTATGGATAATATTAGAAGAGGAGTACCTAAAGGAGTTGAAGTAAGTAATGAA 438

Qy    526 GluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGlu 545
Db    439 CAAATTTATGCTGCTGCTAAAATGGCAAATGCACATGACTTTATTTCAGCAATGCCAGAA 498

Qy    546 LysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArg 565
Db    499 GGATATAACACAATGGTAGGTGATAGAGGTGCACAAATTTCAGGAGGACAAAAACAAAGA 558

Qy    566 LeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThr 585
Db    559 ATTGCTATTGCACGTGCATTGATTAGAAATCCAAAAGTGTTATTACTCGATGAAGCTACA 618

Qy    586 SerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThr 605
Db    619 TCAGCACTTGATTCAGAAAGTGAAAAGATTGTACAAGATGCACTTGACAAAGCAGCAAAA 678

Qy    606 GlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIle 625
Db    679 GGAAGAACAACAATTGTAATTGCACATAGATTATCAACTATTCAAAATGCAGATCAAATA 738

Qy    626 ValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsn 645
Db    739 TGTGTTATTATGAGAGGAAGAATTGCAGAAAGAGGAACACATCAAGAGTTATTAGATTTG 798

Qy    646 ArgAspIleTyrPheLysLeuValAsnAlaGln 656
Db    799 AAAGGATTTTATTATACACTTGCTATGCAACAA 831
```

RESULT 8
AZ682350

```
LOCUS       AZ682350                 871 bp    DNA     linear   GSS 14-DEC-2000
DEFINITION  ENTKB16TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   AZ682350
VERSION     AZ682350.1  GI:11819496
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
  ORGANISM  Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 871)
  AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Forward
            Class: shotgun
            High quality sequence start: 16
            High quality sequence stop: 860.
FEATURES             Location/Qualifiers
     source          1. .871
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a  method described by Clark and Diamond (Clark,
                     C.G., and Diamond, L.S.  (1993) Entamoeba histolytica: a
                     method for isolate identification.  Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + i method used for
                     the library  construction is described in detail in Smith,
                     H.O. and Venter, J.C.  (Making small insert libraries for
                     whole genome shotgun sequencing  projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin  and B.
                     Barell, Oxford University Press, 1999)."
BASE COUNT      349 a    117 c    174 g    231 t
ORIGIN

Alignment Scores:
Pred. No.:              4.83e-78      Length:        871
Score:                  740.50        Matches:       147
Percent Similarity:     70.04%        Conservative:  47
Best Local Similarity:  53.07%        Mismatches:    82
Query Match:            22.28%        Indels:        1
DB:                     17            Gaps:          1

US-09-873-409-1 (1-659) x AZ682350 (1-871)

Qy    381 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 400
Db     39 AATAGTGCTAAAGTTGCCGCATTTAATGTTTATCAAACAATTGATAGAATCCCAGATATT 98

Qy    401 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 420
Db     99 GATTGTCAGTCTATTGGAGGTGAATGTCCAACTGAGTGTAATGGAAATATTAGATTTGAG 158

Qy    421 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 440
Db    159 GATGTTCAATTTGTTTATCCAACAAGACTGTCTCATCATGTATTAAAAGGACTTGACCTT 218

Qy    441 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 460
Db    219 GAAATTAAGAAAGGACAAACAATTGCATTAGTTGGAGCATCAGGATGTGGGAAGTCAACT 278

Qy    461 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 480
Db    279 ACTATTCAATTAATCCAAAGAAATTATGATCCAAATGGTGGAAGAGTAACATTAGACGGA 338

Qy    481 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 500
Db    339 AAAGATATACGAGAGTTGAATATCAAATGGTTAAGAAATCAAATAGGATTAGTTGGACAA 398

Qy    501 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 520
Db    399 GAACCAGTGTTGTTTGCAGGAACAATTCGAGAAAATATTATGCTTGGAGCTAAAGAAGGA 458

Qy    521 ValValProLeu---AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 539
```

```
Db     459 GCAACACCAAGTGAAGAAGAGATGATTGAATGTGCTAAAATGGCAAATGCACATGACTTC 518

Qy     540 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 559
           |||      ||||||||  |||:::|||   :::|||  ||||||||  ||||||
Db     519 ATTTCTAAACTTCCAGAAGGATATGACACAATAATTGGAGAAAAAGGAGCATTATTATCA 578

Qy     560 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 579
           |||||||||||||||||:::||||||||||||||||||::::::  |||  ||||||
Db     579 GGAGGACAAAAACAAAGAATTGCAATTGCACGTGCATTGATTCGAAACCCATCTATTCTT 638

Qy     580 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 599
           |||||||||||||||||||||||||||||         ||||||||:::||||||
Db     639 CTTCTTGATGAAGCTACATCAGCACTTGATACACAAAGTGAAAAGATTGTACAAGAAGCA 698

Qy     600 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 619
           |||:::||||||        ||||||||  ::::::|||  |||||||||:::   :::
Db     699 CTTGAAAAAGCATCTAAAGGAAGAACAACAATTATTGTAGCACATAGACTAACAACTGTT 758

Qy     620 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 639
           :::||||||||  |||   |||   |||  |||:::|||  ||||||||   |||
Db     759 AGAAATGCAGATAAAATATGTGTATTTCATCAAGGAGAAATTATTGAACAAGGAAAACAT 818

Qy     640 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGln 656
           |||||||||:::      :::      |||:::   ||||||    |||
Db     819 CAAGAATTAATGGATTTGAAAGGAACATATTATGGATTAGTCAAAAGACAA 869
```

```
RESULT 9
AZ670821/c
LOCUS       AZ670821                 932 bp    DNA     linear   GSS 14-DEC-2000
DEFINITION  ENTJN69TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   AZ670821
VERSION     AZ670821.1  GI:11807967
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
  ORGANISM  Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 932)
  AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Forward
            Class: shotgun
            High quality sequence start: 46
            High quality sequence stop: 867.
FEATURES             Location/Qualifiers
     source          1..932
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a  method described by Clark and Diamond (Clark,
                     C.G., and Diamond, L.S.  (1993) Entamoeba histolytica: a
                     method for isolate identification.  Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + i method used for
                     the library  construction is described in detail in Smith,
                     H.O. and Venter, J.C.  (Making small insert libraries for
                     whole genome shotgun sequencing  projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin  and B.
                     Barell, Oxford University Press, 1999)."
BASE COUNT      258 a    182 c    125 g    367 t
ORIGIN

Alignment Scores:
Pred. No.:              2.86e-77        Length:         932
Score:                  734.50          Matches:        150
Percent Similarity:     71.11%          Conservative:   42
Best Local Similarity:  55.56%          Mismatches:     77
Query Match:            22.10%          Indels:         2
DB:                     17              Gaps:           1

US-09-873-409-1 (1-659) x AZ670821 (1-932)

Qy     380 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 399
           ::::::  ||||||:::   |||      ::::::     :::::::::   |||:::
Db     835 TTTGCTACAGCTAAAGCTTCTGCATATAGAATTTATCAAACAATTGATAGAATCCCAGAT 776

Qy     400 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 419
           ||||||   ||||||      |||:::   |||      |||   ||||||:::
Db     775 ATTGATTGTAGATCTACAGCTGGTGAATGTCCAACTGAGTGTAATGGAAATATTACATTA 716

Qy     420 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 439
              :::|||   |||   ||||||   ||||||          ||||||   ||||||
Db     715 GAAGATGTTCAATTTAGATATCCAACAAGACCAACTAAACAAATTCTTGGTGGACTTGAT 656

Qy     440 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 459
           |||   |||::::::|||:::|||||||||      ||||||:::||||||||||||||||
Db     655 CTTGAAATTAAGAAAGGACAAACAGTTGCATTAGTAGGAGCATCAGGATGTGGTAAATCA 596

Qy     460 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 479
           |||::::::|||||||:::|||||||   |||||||||||||   |||   |||      |||
Db     595 ACTACTATTCAATTAGTCCAAAGAAATTATGATCCAGTTGGTGGGTCAGTCAAATTAGAT 536

Qy     480 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 499
           |||   |||   ::::::||||||::::::||||||||||:::||||||   :::|||
Db     535 GGAAAAGATTTAAGAGATTTAAATATCAAATGGTTAAGAAATCAAATAGGATTAGTTGGA 476

Qy     500 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 519
           |||||||||:::||||||   |||:::|||   |||||||||         |||
Db     475 CAAGAACCTATTTTGTTTGCATGTACTATTAGAGAAAATATTATGCTTGGAGCTAGAGAT 416

Qy     520 ArgValValProLeu---AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSer 538
                    |||       :::|||:::   |||   |||      ||||||   |||
Db     415 GGAGAAACACCAACTGAAGAAGAGATGATTGAATGTGCTAAAATGGCAAATGCACATGAA 356

Qy     539 PheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeu 558
           ||||||      |||||||||  |||:::|||   ||||||   |||||||||   |||
Db     355 TTTATTTCTCATCTTCCAGAAGGATATGATACAATGGTAGGAGAAAAAGGAGCTGCATTA 296

Qy     559 SerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIle 578
           ||||||||||||||||||||:::||||||||||||||||||::::::||||||   |||
Db     295 TCAGGAGGACAAAAACAAAGAATTGCTATTGCACGTGCATTGATTAGAAAACCTACAATT 236

Qy     579 LeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHis 598
           |||||||||||||||||||||||||||||||         |||||||||:::||||||
Db     235 TTATTACTTGATGAAGCTACATCAGCACTTGATACACAAAGTGAAAAGATTGTACAACAA 176

Qy     599 AlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAla 618
           ||||||:::||||||      |||||||||   :::||||||   |||||||||:::
Db     175 GCACTTGAAAAAGCTAGTCAAGGAAGAACAACAATTGTTGTAGCACATAGATTAACAACT 116

Qy     619 IleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThr 638
           ::::::||||||      |||   |||   |||   |||:::|||   ||||||||||||
Db     115 GTTAGAAATGCAAGTAGAATTTGTGTATTCCATCAAGGAGAAATTAT-GAACAAGGAACA 57

Qy     639 HisGlnGluLeuLeuArgAsnArgAspIle 648
           ||||||||||||         :::|||:::
Db      56 CATCAAGAATTAAGGATTTAAAAGGATTTA 27

RESULT 10
AZ683753
LOCUS       AZ683753                 947 bp    DNA     linear   GSS 14-DEC-2000
```

```
DEFINITION  ENTIL96TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   AZ683753
VERSION     AZ683753.1  GI:11820899
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
  ORGANISM  Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 947)
  AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Forward
            Class: shotgun
            High quality sequence start: 40
            High quality sequence stop: 778.
FEATURES             Location/Qualifiers
     source          1. .947
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a  method described by Clark and Diamond (Clark,
                     C.G., and Diamond, L.S.  (1993) Entamoeba histolytica: a
                     method for isolate identification.  Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + i method used for
                     the library  construction is described in detail in Smith,
                     H.O. and Venter, J.C.  (Making small insert libraries for
                     whole genome shotgun sequencing  projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin  and B.
                     Barell, Oxford University Press, 1999)."
BASE COUNT      379 a    121 c    178 g    269 t
ORIGIN

Alignment Scores:
Pred. No.:               1.55e-76       Length:        947
Score:                   728.50         Matches:       149
Percent Similarity:      69.82%         Conservative:  43
Best Local Similarity:   54.18%         Mismatches:    81
Query Match:             21.92%         Indels:        2
DB:                      17             Gaps:          1

US-09-873-409-1 (1-659) x AZ683753 (1-947)

Qy    383 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSer 402
           ||||||:::    |||        :::|||        :::::::::    |||:::||||||
Db     69 GCAAAAGCATCAGCCTATAAAATATTTACAACAATTGATAGAATCCCAGATATTGATTGT 128

Qy    403 ArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 422
           :::|||        |||:::    |||          |||   ||||||:::    |||    :::|||
Db    129 CAGTCTATTGGAGGTGAATGTCCAACTGAGTGTAATGGGAATATTAGATTTGATGATGTT 188

Qy    423 SerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIle 442
              |||    ||||||    ||||||              :::|||:::|||:::    |||    |||
Db    189 CAATTTGTTTATCCAACAAGACCATCTCATCATGTATTAAAAGGAATTGACCTTGAAATT 248

Qy    443 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 462
          ::::::|||:::    :::|||    ||||||:::|||||||||    |||||||||::::::
Db    249 AAGAAAGGAGAAATAATTGCATTAGTTGGAGCATCAGGATGTAGAAAGTCAACTACTATT 308

Qy    463 GlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAsp 482
           ||||||:::||||||    |||:::|||          |||:::|||          ||||||   |||
Db    309 CAATTAATCCAAAGAAATTATGAACCAAATGGTGGAAGAGTAACATTAGATGGGAAAGAC 368

Qy    483 AlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 502
             :::|||||||||:::::::|||||||||:::||||||      ||||||    |||||||||
Db    369 ATAAGAGAGTTGAATATCAAATGGTTAAGAAATCAAATAGGAATAGTTGGACAAGAACCA 428

Qy    503 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGly---AspAsnSerArgVal 521
           |||||||||:::    :::|||    |||||||||          |||
Db    429 GTATTATTTTCAGGAACAATTGGAGAAAATATCATGCTTGGAGCTAAAGAAGGAGAAACA 488

Qy    522 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 541
           :::         :::|||:::    |||                  ||||||    |||    ||||||
Db    489 CTAAGTAAAGAAGAGATGATTGAATGTACTAAAATGGCAAATGCACATGACTTTATTTCT 548

Qy    542 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 561
              |||||||||    |||:::|||      :::|||      |||||||||    ||||||||||||
Db    549 AAACTTCCAGAAGGATATGACACAATAATTGGAGAAAAAGGAGCATTATTATCAGGAGGA 608

Qy    562 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 581
           ||||||||||||:::||||||||||||||||||::::::    |||    ||||||||||||
Db    609 CAAAAACAAAGAATTGCAATTGCACGTGCATTGATTCGAAACCCATCTATTCTTCTTCTT 668

Qy    582 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 601
           |||||||||||||||||||||||||          |||||||||:::||||||    |||||||||
Db    669 GATGAAGCTACATCAGCACTTGATACACAAAGTGAAAAGATTGTACAAGAAGCACTTGAC 728

Qy    602 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 621
           ||||||      |||||||||    ::::::|||    ||||||||||||    |||||||||
Db    729 AAAGCATCAAAAGGAAGAACAACAATTATTGTAGGACATAGACTATCGACTATTCAAAAT 788

Qy    622 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 641
           ||||||    |||    |||:::          |||||||||    ||||||||||||||||||||
Db    789 GCAGATCAAATATGTGTTATTATGAGAGGAAAAATAGTAGAACAAGGAACACATCAAGAA 848

Qy    642 -LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 655
            ||||||          :::      |||:::    |||        |||
Db    849 TTTATTGGATTTGAAAAGATTTATTATACACTTGCTTATGCA 891

RESULT 11
AZ540627
LOCUS       AZ540627                 886 bp    DNA     linear   GSS 14-NOV-2000
DEFINITION  ENTEQ18TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   AZ540627
VERSION     AZ540627.1  GI:11147603
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
  ORGANISM  Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 886)
  AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Forward
            Class: shotgun
            High quality sequence start: 23
            High quality sequence stop: 856.
FEATURES             Location/Qualifiers
     source          1. .886
                     /organism="Entamoeba histolytica"
```

```
                /strain="HM1:IMSS"
                /db_xref="taxon:5759"
                /clone_lib="Entamoeba histolytica Sheared DNA"
                /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                Institute for Genomic Research (TIGR), Rockville, MD.
                Genomic DNA isolated from broth cultures of E. histolytica
                using a  method described by Clark and Diamond (Clark,
                C.G., and Diamond, L.S.  (1993) Entamoeba histolytica: a
                method for isolate identification.  Exp. Parasitol.
                77:450.). The DNA was mechanically sheared to give a
                tight size distribution (~2 kb). The v + i method used for
                the library  construction is described in detail in Smith,
                H.O. and Venter, J.C.  (Making small insert libraries for
                whole genome shotgun sequencing  projects. In Genome
                Sequencing: A Practical Approach, eds. M. Vaudin  and B.
                Barell, Oxford University Press, 1999)."
BASE COUNT      364 a    107 c    181 g    234 t
ORIGIN

Alignment Scores:
Pred. No.:              2.42e-76       Length:          886
Score:                  726.50         Matches:         149
Percent Similarity:     72.08%         Conservative:    42
Best Local Similarity:  56.23%         Mismatches:      72
Query Match:            21.86%         Indels:          3
DB:                     17             Gaps:            1

US-09-873-409-1 (1-659) x AZ540627 (1-886)

Qy    394 LeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCys 413
          :::::::::   |||:::|||||   ||||||        |||        |||    |||
Db     22 ATTGATAGAATACCAGATATTGATTGTAGATCTACAGCTGGATTAGTTCCAACTGAGTGT 81

Qy    414 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 433
             ||||||:::::          :::|||    |||    ||||||    ||||||:::
Db     82 ATTGGAAATATTAAATTAGAAGATGTTCAATTTAGATATCCAACAAGACCAAATAAACAA 141

Qy    434 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 453
          ||||||   ||||||   |||   |||::::::|||:::|||||||||   ||||||:::
Db    142 ATTCTTGGTGGACTTGACCTTGAAATTAAGAAAGGAGAAACAGTTGCATTAGTAGGAGCA 201

Qy    454 SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGln 473
          |||||||||||||||||||||||||::::::||||||:::||||||:::||||||||||||
Db    202 TCAGGATGTGGTAAGTCAACTACAATTCAATTAGTCCAAAGAGTCTATGATCCAGTTGGT 261

Qy    474 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 493
          |||:::|||         ||||||   |||   :::|||||||||::::::|||||||||:::
Db    262 GGAAAAGTAACATTAGATGGAAATGATTTAAGAGAGTTGAACTTGAAATGGTTAAGAAAT 321

Qy    494 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 513
          ||||||   :::|||   |||||||||:::||||||   |||:::|||   |||||||||
Db    322 CAAATAGGATTAGTTGGACAAGAACCTATTTTGTTTGCATGTACTATTAGAGAAAATATT 381

Qy    514 AlaTyrGlyAspAsnSerArgValValProLeu---AspGluIleLysGluAlaAlaAsn 532
                |||                 |||      :::|||:::   |||   |||
Db    382 ATGCTTGGAGCTAAAGATGGAGAAACACCAACTGAAGAAGAGATGATTGAATGTGCTAAA 441

Qy    533 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 552
             ||||||   |||   ||||||         |||||||||   |||:::|||   ||||||
Db    442 ATGGCAAATGCACATGAATTTATTTCTCATCTTCCAGAAGGATATGATACAATGGTAGGA 501

Qy    553 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 572
             :::||||||   ||||||||||||||||||||||||:::||||||||||||||||||
Db    502 GAAAGAGGAGCTGCATTATCAGGAGGACAAAAACAAAGAATTGCTATTGCACGTGCATTG 561

Qy    573 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 592
          ::::::||||||   ||||||||||||||||||||||||||||||||||||      |||
Db    562 ATTAGAAAACCTACAATTTTATTACTTGATGAAGCTACATCAGCACTTGATACACAAAGT 621

Qy    593 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 612
          ||||||:::||||||   ||||||:::||||||         |||||||||   ::::::|||
Db    622 GAAAAGATTGTACAACAAGCACTTGAAAAAGCTAGTCAAGGAAGAACAACAATTATTGTA 681

Qy    613 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeu-IleValValLeuHisAsnGlyLy 632
          |||||||||:::   ::::::|||||||||:::   |||   |||   |||   |||::
Db    682 GCACATAGATTAACAACTGTTAGAAATGCAAATAGAAATTTGTGTATTCCATCAAGGAGA 741

Qy    632 sIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLe 652
           :|||   ||||||||||||||||||||||||:::                |||:::   ||
Db    742 AATTATTGAACAAGGAACACATCAAGAATTAATGGATTT-GAAAGCACATATTATGGATT 800

Qy    652 uValAsnAlaGln 656
           ||||      |||
Db    801 AGTTAAAAGACAA 813

RESULT 12
AZ687805/c
LOCUS       AZ687805                 880 bp    DNA     linear   GSS 14-DEC-2000
DEFINITION  ENTLU52TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   AZ687805
VERSION     AZ687805.1  GI:11824951
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
  ORGANISM  Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 880)
  AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Forward
            Class: shotgun
            High quality sequence start: 30
            High quality sequence stop: 858.
FEATURES             Location/Qualifiers
     source          1..880
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a  method described by Clark and Diamond (Clark,
                     C.G., and Diamond, L.S.  (1993) Entamoeba histolytica: a
                     method for isolate identification.  Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + i method used for
                     the library  construction is described in detail in Smith,
                     H.O. and Venter, J.C.  (Making small insert libraries for
                     whole genome shotgun sequencing  projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin  and B.
                     Barell, Oxford University Press, 1999)."
BASE COUNT      247 a    167 c    122 g    344 t
ORIGIN

Alignment Scores:
Pred. No.:              5.5e-76        Length:          880
Score:                  723.50         Matches:         151
Percent Similarity:     69.93%         Conservative:    42
Best Local Similarity:  54.71%         Mismatches:      82
Query Match:            21.77%         Indels:          2
DB:                     17             Gaps:            1

US-09-873-409-1 (1-659) x AZ687805 (1-880)
```

```
Qy     383 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSer 402
Db     831 GCTAAAGCTTCTGCATATAGAATTTATCAAACAATTGATAGAATCCCAGATATTGATTGT 772

Qy     403 ArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 422
Db     771 AGATCTACAGCTGGTGAATGTCCAACTGAGTGTAATGGAAATATTACATTAGAAGATGTT 712

Qy     423 SerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIle 442
Db     711 CAATTTAGATATCCAACAAGACCAACTAAACAAATTCTTGGTGGACTTGATCTTGAAATT 652

Qy     443 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 462
Db     651 AAGAAAGGACAAACAGTTGCATTAGTAGGAGCATCAGGATGTGGTAAATCAACTACTATT 592

Qy     463 GlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAsp 482
Db     591 CAATTAGTCCAAAGAAATTATGATCCAGTTGGTGGGTCAGTCAAATTAGATGGAAAAGAT 532

Qy     483 AlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 502
Db     531 TTAAGAGATTTAAATATCAAATGGTTAAGAAATCAAATAGGATTAGTTGGACAAGAACCT 472

Qy     503 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValVal 522
Db     471 ATTTTGTTTGCATGTACTATTAGAGAAAATATTATGCTTGGAGCTAGAGATGGAGAAACA 412

Qy     523 ProLeu---AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 541
Db     411 CCAACTGAAGAAGAGATGATTGAATGTGCTAAAATGGCAAATGCACATGAATTTATTTCT 352

Qy     542 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 561
Db     351 CATCTTCCAGAAGGATATGATACAATGGTAGGAGAAAAAGGAGCTGCATTATCAGGAGGA 292

Qy     562 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 581
Db     291 CAAAAACAAAGAATTGCTATTGCACGTGCATTGATTAGAAAACCTACAATTTTATTACTT 232

Qy     582 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 601
Db     231 GATGAAGCTACATCAGCACTTGATACACAAAGTGAAAAGATTGTACAACAAGCACTTGAA 172

Qy     602 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 621
Db     171 AAAGCTAGTCAAGGAAGAACAACAATTGTTGTAGCACATAGATTAACAACTGTTAGAAAT 112

Qy     622 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 641
Db     111 GCAAGTAGAATTTGTGTATTCCATCAAGGAGAAATTATTAA-CAAGGAACACATCAAGAA 53

Qy     642 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db      52 TTAAGGATTTAAAGCAACATATTATGGATTATTTTTCCAGCACAAACT 5
```

RESULT 13
BQ720763

```
LOCUS       BQ720763             939 bp    mRNA    linear   EST 16-JUL-2002
DEFINITION  AGENCOURT_8233170 Lupski_dorsal_root_ganglion Homo sapiens cDNA
            clone IMAGE:6184557 5', mRNA sequence.
ACCESSION   BQ720763
VERSION     BQ720763.1  GI:21859660
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 939)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. James R. Lupski
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM13573  row: o  column: 22
            High quality sequence stop: 646.
FEATURES             Location/Qualifiers
     source          1. .939
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6184557"
                     /clone_lib="Lupski_dorsal_root_ganglion"
                     /sex="male"
                     /tissue_type="dorsal root ganglia"
                     /dev_stage="adult, 36 yr"
                     /lab_host="DH10B"
                     /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:
                     NotI; Site_2: SalI; cDNA made by oligo-dT priming.
                     Directionally cloned using the following adaptors:
                     5'-TCGACCCACGCGTCCG-3' and
                     5'-GACTAGTTCTAGATCGCGAGCGGCCGCCCT(15)-3'. Size selected >
                     1 kb for average insert length 1.7 kb. This is a primary
                     library, non-amplified. Library constructed by Life
                     Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
                     College of Medicine) and is available through Life
                     Technologies."
BASE COUNT      272 a    202 c    221 g    244 t
ORIGIN

Alignment Scores:
Pred. No.:              1.07e-75        Length:         939
Score:                  721.50          Matches:        140
Percent Similarity:     74.44%          Conservative:   61
Best Local Similarity:  51.85%          Mismatches:     66
Query Match:            21.71%          Indels:         3
DB:                     14              Gaps:           1

US-09-873-409-1 (1-659) x BQ720763 (1-939)

Qy     161 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 180
Db       3 AAAGCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGA 62

Qy     181 GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 200
Db      63 CAGGATGTGAGTTGGTTTGATGAGCCTAAAAACACCACTGGAGCATTGACTACCAGGCTC 122

Qy     201 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 220
Db     123 GCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAG 182

Qy     221 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr 240
Db     183 AATATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACA 242

Qy     241 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 260
Db     243 CTGTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATG 302

Qy     261 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 280
Db     303 TTGTCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGTTCTGGGAAGATCGCTACT 362

Qy     281 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 300
Db     363 GAAGCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACAT 422

Qy     301 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 320
```

```
Db     423 ATGTATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTT 482

Qy     321 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 340
           |||      ::::::|||:::   |||   :::||||||:::||||||     ||||||
Db     483 GGAATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGG 542

Qy     341 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 360
           |||||||||||||||:::              |||:::   |||    :::   :::||||||
Db     543 TTTGGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTT 602

Qy     361 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 380
           :::|||:::   :::||||||||||||:::|||:::                ||||||:::|||
Db     603 TCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTAT 662

Qy     381 Ser-LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIl 400
           ::: |||||||||         |||||||||:::        ::::::||||||   |||   ||
Db     663 GCCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGAT 722

Qy     400 eAspSerArgSer-GlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGlu---P 419
           |||||||   |||      ||||||         |||:::|||   |||||||||:::      :
Db     723 TGACAGCTACAGCACCGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACCATTT 782

Qy     419 heArgGluValSerPhePheTyrPro 427
           ::   ::::::   |||   ||||||
Db     783 GGTGAAAGTTGTATTTCAACTATCCC 808

RESULT 14
AZ541090
LOCUS       AZ541090                897 bp    DNA     linear   GSS 14-NOV-2000
DEFINITION  ENTDS67TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   AZ541090
VERSION     AZ541090.1  GI:11148493
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
  ORGANISM  Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 897)
  AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Reverse
            Class: shotgun
            High quality sequence start: 81
            High quality sequence stop: 890.
FEATURES             Location/Qualifiers
     source          1..897
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a  method described by Clark and Diamond (Clark,
                     C.G., and Diamond, L.S.  (1993) Entamoeba histolytica: a
                     method for isolate identification.  Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + i method used for
                     the library  construction is described in detail in Smith,
                     H.O. and Venter, J.C.  (Making small insert libraries for
                     whole genome shotgun sequencing  projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin  and B.
                     Barell, Oxford University Press, 1999)."
BASE COUNT      335 a    110 c    177 g    275 t
ORIGIN

Alignment Scores:
Pred. No.:              1.31e-75      Length:          897
Score:                  720.50        Matches:         143
Percent Similarity:     74.59%        Conservative:    39
Best Local Similarity:  58.61%        Mismatches:      61
Query Match:            21.68%        Indels:          1
DB:                     17            Gaps:            1

US-09-873-409-1 (1-659) x AZ541090 (1-897)

Qy     414 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 433
           :::|||   :::||||||:::::::::   |||   ||||||   |||||||||
Db      94 AAAGGTGAAATTGAATTTAAAGACATTTGTTTCAGATATCCAACAAGACCAGACAATTCT 153

Qy     434 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 453
           :::|||:::|||:::|||         :::|||:::|||||||||||||||   ||||||:::
Db     154 GTCTTGAAAGGTATTTCATTCAAAGTAGAACAAGGAAAAACTGTTGCATTAGTAGGAGCA 213

Qy     454 SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGln 473
           ||||||||||||||||||||||||||||||||||||:::::::|||   |||||||||
Db     214 TCAGGATGTGGTAAATCAACATCAGTTCAGTTGATTGAAAGATTTTATGATCCAACACAT 273

Qy     474 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 493
           |||   ||||||   ||||||      :::   |||:::||||||:::      :::|||||||||
Db     274 GGAGATGTATTATTAGATGGACATAATATCAAAGATTTGAATATTCATTTCTTAAGAAGT 333

Qy     494 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 513
           ||||||   :::|||   ||||||||||||||||||      |||:::   :::||||||
Db     334 CAAATTGGAATGGTAGGACAAGAACCAGTATTATTTGCTGAAAGTGTTATGGATAATATT 393

Qy     514 AlaTyrGlyAspAsnSerArgVal---ValProLeuAspGluIleLysGluAlaAlaAsn 532
                 |||            |||   |||      ::::::|||      ||||||
Db     394 AGAAGAGGAGTACCTAAAGGAGTTGAAGTAAGTAATGAACAAATTTATGCTGCTGCTAAA 453

Qy     533 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 552
              ||||||   |||   ||||||      :::||||||   |||||||||   ||||||
Db     454 ATGGCAAATGCACATGACTTTATTTCAGCAATGCCAGAAGGATATAACACAATGGTAGGT 513

Qy     553 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 572
           :::|||||||||:::|||||||||||||||||||||||||||||||||||||||||||
Db     514 GATAGAGGTGCACAAATTTCAGGAGGACAAAAACAAAGAATTGCTATTGCACGTGCATTG 573

Qy     573 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 592
           ::::::   ||||||:::||||||||||||||||||||||||||||||||||::::::|||
Db     574 ATTAGAAATCCAAAAGTGTTATTACTCGATGAAGCTACATCAGCACTTGATTCAGAAAGT 633

Qy     593 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 612
           ||||||:::||||||   ||||||||||||||||      |||||||||   :::|||:::
Db     634 GAAAAGATTGTACAAGATGCACTTGACAAAGCAGCAAAAGGAAGAACAACAATTGTAATT 693

Qy     613 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 632
              ||||||||||||   ||||||||||||||||   |||   |||:::      |||:::
Db     694 GCACATAGATTATCAACTATTCAAAATGCAGATCAAATATGTGTTATTATGAGAGGAAGA 753

Qy     633 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 652
           |||   |||:::|||||||||||||||||||||      :::      |||:::   |||
Db     754 ATTGCAGAAAGAGGAACACATCAAGAGTTATTAGATTTGAAAGGATTTTATTATACACTT 813

Qy     653 ValAsnAlaGln 656
                 |||
Db     814 GCTATGCAACAA 825

RESULT 15
BH155700
LOCUS       BH155700                913 bp    DNA     linear   GSS 24-SEP-2001
DEFINITION  ENTRO54TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   BH155700
VERSION     BH155700.1  GI:15727822
```

```
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
  ORGANISM  Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 913)
  AUTHORS   Loftus,B., Wang,Z., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library (2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Reverse
            Class: shotgun
            High quality sequence start: 7
            High quality sequence stop: 805.
FEATURES             Location/Qualifiers
     source          1. .913
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a  method described by Clark and Diamond (Clark,
                     C.G., and Diamond, L.S.  (1993) Entamoeba histolytica: a
                     method for isolate identification.  Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + i method used for
                     the library  construction is described in detail in Smith,
                     H.O. and Venter, J.C.  (Making small insert libraries for
                     whole genome shotgun sequencing  projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin  and B.
                     Barell, Oxford University Press, 1999)."
BASE COUNT     359 a    118 c    185 g    251 t
ORIGIN

Alignment Scores:
Pred. No.:              1.64e-74        Length:         913
Score:                  711.50          Matches:        146
Percent Similarity:     69.37%          Conservative:   42
Best Local Similarity:  53.87%          Mismatches:     82
Query Match:            21.41%          Indels:         2
DB:                     17              Gaps:           1

US-09-873-409-1 (1-659) x BH155700 (1-913)

Qy     380 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 399
           ::::::   ||||||:::   |||        ::::::      :::::::::   |||:::
Db      29 TTTGCTACAGCTAAAGCTTCTGCATATAGAATTTATCAAACAATTGATAGAATCCCAGAT 88

Qy     400 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 419
           ||||||   ||||||          |||:::   |||      |||   ||||||:::
Db      89 ATTGATTGTAGATCTACAGCTGGTGAATGTCCAACTGAGTGTAATGGAAATATTACATTA 148

Qy     420 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 439
              :::|||   |||   ||||||   ||||||      ||||||   ||||||
Db     149 GAAGATGTTCAATTTAGATATCCAACAAGACCAACTAAACAAATTCTTGGTGGACTTGAT 208

Qy     440 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 459
           |||   |||::::::|||:::|||||||||   ||||||:::||||||||||||||||||
Db     209 CTTGAAATTAAGAAAGGACAAACAGTTGCATTAGTAGGAGCATCAGGATGTGGTAAATCA 268

Qy     460 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 479
           |||::::::||||||:::||||||   ||||||||||||   |||   |||      |||
Db     269 ACTACTATTCAATTAGTCCAAAGAAATTATGATCCAGTTGGTGGGTCAGTCAAATTAGAT 328

Qy     480 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 499
           |||   |||   ::::::||||||::::::|||||||||:::||||||      :::|||
Db     329 GGAAAAGATTTAAGAGATTTAAATATCAAATGGTTAAGAAATCAAATAGGATTAGTTGGA 388

Qy     500 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 519
           |||||||||:::||||||   |||:::|||   |||||||||         |||
Db     389 CAAGAACCTATTTTGTTTGCATGTACTATTAGAGAAAATATTATGCTTGGAGCTAGAGAT 448

Qy     520 ArgValValProLeu---AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSer 538
                       |||      :::|||:::   |||   |||      ||||||   |||
Db     449 GGAGAAACACCAACTGAAGAAGAGATGATTGAATGTGCTAAAATGGCAAATGCACATGAA 508

Qy     539 PheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeu 558
           ||||||         |||||||||   |||:::|||   ||||||   |||||||||   |||
Db     509 TTTATTTCTCATCTTCCAGAAGGATATGATACAATGGTAGGAGAAAAAGGAGCTGCATTA 568

Qy     559 SerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIle 578
           ||||||||||||||||||:::||||||||||||||||||::::::||||||      |||
Db     569 TCAGGAGGACAAAAACAAAGAATTGCTATTGCACGTGCATTGATTAGAAAACCTACAATT 628

Qy     579 LeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHis 598
           ||||||||||||||||||||||||||||||         |||||||||:::||||||
Db     629 TTATTACTTGATGAAGCTACATCAGCACTTGATACACAAAGTGAAAAGATTGTACAACAA 688

Qy     599 AlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAla 618
           ||||||:::||||||      |||||||||      :::||||||   |||||||||:::
Db     689 GCACTTGAAAAAGCTAGTCAAGGAAGAACAACAATTGTTGTAGCACATAGATTAACAACT 748

Qy     619 IleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThr 638
           ::::::||||||            |||   |||   |||:::|||         ||||||
Db     749 GTTAGAAATGCAAGTAGAAATTGTGTATTCCATCAAGGAGAAATTAT-TTGACAGGAACA 807

Qy     639 HisGlnGluLeuLeuArgAsnArgAspIleTyr 649
           |||   ||||||:::         :::   |||
Db     808 CATTCAGAATTAATGGATTTGAAGGGGACATAT 840

Search completed: March 31, 2003, 13:49:30
Job time : 4561.85 secs
```

```
                    GenCore version 5.1.4_p5_4578
           Copyright (c) 1993 - 2003  Compugen Ltd.


OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 31, 2003, 02:24:20 ; Search time 96.5072 Seconds
                                          (without alignments)
                                          2094.145 Million cell updates/sec

Title:          US-09-873-409-1
Perfect score:  3323
Sequence:       1 MLAEKGAHAELMAKRGLYYS..........QELLRNRDIYFKLVNAQSVQ 659

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters:      882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115422_19259/app_query.fasta_1.7544
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_578_@runat_27032003_115422_19259 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Issued_Patents_NA:*
                1:  /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
                2:  /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
                3:  /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
                4:  /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
                5:  /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
                6:  /cgn2_6/ptodata/2/ina/backfiles1.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1872	56.3	2726	1	US-08-461-823-1	Sequence 1, Appli
2	1872	56.3	4264	2	US-08-784-649A-1	Sequence 1, Appli
3	1872	56.3	4264	2	US-08-784-649A-5	Sequence 5, Appli
4	1872	56.3	4646	1	US-08-181-471-2	Sequence 2, Appli
5	1872	56.3	4669	6	5206352-3	Patent No. 5206352
6	1872	56.3	6505	2	US-08-793-610-5	Sequence 5, Appli
7	1872	56.3	9318	2	US-08-793-610-6	Sequence 6, Appli
8	1855	55.8	4669	2	US-08-752-447-1	Sequence 1, Appli
9	1855	55.8	4669	4	US-09-316-167-1	Sequence 1, Appli
10	1849.5	55.7	4233	3	US-09-120-513-1	Sequence 1, Appli
11	1849.5	55.7	4233	4	US-09-450-105-1	Sequence 1, Appli
12	1845	55.5	4669	2	US-08-583-276-18	Sequence 18, Appl
13	1254	37.7	4002	2	US-08-996-545-1	Sequence 1, Appli
14	1254	37.7	4002	2	US-08-996-545-3	Sequence 3, Appli
15	1254	37.7	4002	4	US-09-328-320-1	Sequence 1, Appli
16	1254	37.7	4002	4	US-09-328-320-3	Sequence 3, Appli
17	1227.5	36.9	4047	2	US-08-612-734B-1	Sequence 1, Appli
18	1184.5	35.6	4800	2	US-08-612-734B-3	Sequence 3, Appli
19	1151	34.6	3924	1	US-08-395-246C-1	Sequence 1, Appli
20	1049.5	31.6	4224	1	US-08-612-521-1	Sequence 1, Appli
21	975	29.3	3909	1	US-08-232-537-1	Sequence 1, Appli
22	851	25.6	6143	1	US-08-612-521-3	Sequence 3, Appli
23	842	25.3	3924	2	US-08-996-644-3	Sequence 3, Appli
24	842	25.3	3924	3	US-09-352-552-3	Sequence 3, Appli
25	842	25.3	3927	2	US-08-996-644-1	Sequence 1, Appli
26	842	25.3	3927	3	US-09-352-552-1	Sequence 1, Appli
27	840	25.3	3792	4	US-09-351-224E-10	Sequence 10, Appl
28	783	23.6	3999	4	US-09-351-224E-9	Sequence 9, Appli
29	677	20.4	2376	1	US-08-394-880B-1	Sequence 1, Appli
30	644	19.4	2244	4	US-09-061-764A-18	Sequence 18, Appl
31	623.5	18.8	2061	4	US-09-061-764A-17	Sequence 17, Appl
32	608.5	18.3	5120	3	US-08-772-270A-6	Sequence 6, Appli
33	608.5	18.3	8370	2	US-08-488-706-1	Sequence 1, Appli
34	601	18.1	1749	4	US-09-134-001C-1893	Sequence 1893, Ap
35	600.5	18.1	1743	4	US-09-134-001C-2774	Sequence 2774, Ap
36	599	18.0	13188	4	US-08-961-527-70	Sequence 70, Appl
37	594.5	17.9	7721	3	US-08-772-270A-14	Sequence 14, Appl
38	555.5	16.7	1959	4	US-09-061-764A-4	Sequence 4, Appli
39	555.5	16.7	7760	4	US-08-961-527-63	Sequence 63, Appl
c 40	552.5	16.6	28804	2	US-08-592-874-1	Sequence 1, Appli
c 41	552.5	16.6	28804	3	US-09-096-942-2	Sequence 2, Appli
c 42	552.5	16.6	28804	3	US-09-096-867-2	Sequence 2, Appli
43	548	16.5	6492	4	US-08-961-527-188	Sequence 188, App
44	544.5	16.4	4411529	4	US-09-103-840A-1	Sequence 1, Appli
45	543.5	16.4	7186	4	US-08-961-527-39	Sequence 39, Appl

ALIGNMENTS

```
RESULT 1
US-08-461-823-1
; Sequence 1, Application US/08461823
; Patent No. 5593840
;  GENERAL INFORMATION:
;    APPLICANT:  Bhatnagar, Satish K.
;    APPLICANT:  George Jr., Albert L.
;    APPLICANT:  Nazarenko, Irina
;    TITLE OF INVENTION:  AMPLIFICATION OF NUCLEIC ACID SEQUENCES
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OncorPharm, Inc.
;      STREET:  200 Perry Parkway
;      CITY:  Gaithersburg
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20877
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/461,823
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/168,621
;      FILING DATE:  16-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/010,433
;      FILING DATE:  27-JAN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Karta, Glenn E.
;      REGISTRATION NUMBER:  30,649
;      REFERENCE/DOCKET NUMBER:  PA-0012 CIP 2
```

```
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301 527-2058
;      TELEFAX:  301 208-6997
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2726 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-461-823-1

Alignment Scores:
Pred. No.:               2.8e-219      Length:         2726
Score:                   1872.00       Matches:        356
Percent Similarity:      74.22%        Conservative:   142
Best Local Similarity:   53.06%        Mismatches:     159
Query Match:             56.33%        Indels:         14
DB:                      1             Gaps:           4

US-09-873-409-1 (1-659) x US-08-461-823-1 (1-2726)

Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db    317 GTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAA 376

Qy     21 LeuValMetSerGln------------------AspIleLysLysAlaAspGluGlnMet 34
Db    377 CTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCC 436

Qy     35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
Db    437 AAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGA 496

Qy     54 ---------------LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
Db    497 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACC 556

Qy     68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db    557 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAAT 616

Qy     88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db    617 TTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTG 676

Qy    108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
Db    677 CAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGAT 736

Qy    127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db    737 CCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATT 796

Qy    147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db    797 ATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTC 856

Qy    167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
Db    857 ACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 916

Qy    187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db    917 GATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAA 976

Qy    207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db    977 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGG 1036

Qy    227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
Db   1037 ACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATT 1096

Qy    247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db   1097 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 1156

Qy    267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db   1157 AAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 1216

Qy    287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db   1217 CGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTG 1276

Qy    307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db   1277 CAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTC 1336

Qy    327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
Db   1337 ACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTG 1396

Qy    347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db   1397 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGT 1456

Qy    367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db   1457 GCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCA 1516

Qy    387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Db   1517 GCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAA 1576

Qy    407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
Db   1577 GGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTAT 1636

Qy    427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db   1637 CCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAG 1696

Qy    447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
Db   1697 ACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAG 1756

Qy    467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db   1757 CGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTG 1816

Qy    487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db   1817 AATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGAC 1876

Qy    507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db   1877 TGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAG 1936

Qy    527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db   1937 ATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAA 1996

Qy    547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Db   1997 TATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATT 2056

Qy    567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Db   2057 GCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCA 2116

Qy    587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
```

```
Db     2117 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 2176

Qy      607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
            ||||||||:::|||:::    |||||||||||    ||||||||||||||||||||||
Db     2177 CGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTG 2236

Qy      627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
              |||     ||||||:::::||||||   ||||||||||||:::||||||        :::
Db     2237 GTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAA 2296

Qy      647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
                   |||||||||   :::|||:::   |||:::
Db     2297 GGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 2329
```

RESULT 2
US-08-784-649A-1
; Sequence 1, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
; APPLICANT: Sikic, Branimir I
; APPLICANT: Chen, Gang
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
; TITLE OF INVENTION: CYCLOSPORIN MODULATION
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/784,649A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: Reg.No. 5830697 36,677
; REFERENCE/DOCKET NUMBER: 06037/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-784-649A-1

Alignment Scores:

Pred. No.:	6.17e-219	Length:	4264
Score:	1872.00	Matches:	356
Percent Similarity:	74.22%	Conservative:	142
Best Local Similarity:	53.06%	Mismatches:	159
Query Match:	56.33%	Indels:	14
DB:	2	Gaps:	4

US-09-873-409-1 (1-659) x US-08-784-649A-1 (1-4264)

```
Qy        1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
            ::::::    ||||||||    |||    ||||||||    ::::::|||:::|||:::
Db     1951 GTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAA 2010

Qy       21 LeuValMetSerGln------------------AspIleLysLysAlaAspGluGlnMet 34
            |||||      |||                    :::::::::   |||   ::::::
Db     2011 CTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCC 2070

Qy       35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
            :::|||          :::   |||      :::|||                 :::
Db     2071 AAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGA 2130

Qy       54 ---------------LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
                           :::|||::::::   |||         |||:::         :::
Db     2131 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACC 2190

Qy       68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
              ::::::   :::   |||:::|||   ||||||      :::|||:::||||||||
Db     2191 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAAT 2250

Qy       88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
                  |||||||||:::   |||:::|||         ::::::::||||||   :::
Db     2251 TTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTG 2310

Qy      108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
              |||   |||:::|||||||||:::|||||||||      :::|||         :::|||
Db     2311 CAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGAT 2370

Qy      127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
                  |||   :::   ::::::   ::::::|||::::::|||:::   ||||||:::
Db     2371 CCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATT 2430

Qy      147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
            |||   |||:::::::::|||:::||||||         :::|||:::||||||||||||||
Db     2431 ATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTC 2490

Qy      167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
            |||   |||||||||::::::   |||::::::||||||   ||||||::::::||||||
Db     2491 ACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 2550

Qy      187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
            |||:::   :::|||:::||||||   |||||||||   ||||||   |||   ||||||
Db     2551 GATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAA 2610

Qy      207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
            ::::::||||||   |||||||||:::   |||:::|||||||||      |||:::|||
Db     2611 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGG 2670

Qy      227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
               :::|||||||||||||||||||||||||||::::::|||   |||:::|||:::|||
Db     2671 ACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATT 2730

Qy      247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
               |||:::::|||:::   |||::::::|||         ::::::|||   |||
Db     2731 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 2790

Qy      267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
            |||||||||:::||||||:::   |||||||||||||||||||||||||||:::||||||
Db     2791 AAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 2850

Qy      287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
            ||||||:::||||||||||||:::|||:::   ||||||   ||||||   :::   |||
Db     2851 CGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTG 2910

Qy      307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
            |||      :::||||||:::   :::||||||   |||   |||      ::::::|||
Db     2911 CAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTC 2970

Qy      327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
            :::   |||   :::||||||:::|||   ||||||||||||||||||||||||:::
Db     2971 ACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTG 3030

Qy      347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
                  |||:::   |||   :::   :::||||||:::|||:::      :::|||
Db     3031 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGT 3090

Qy      367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
            |||||||||:::|||:::               ||||||:::|||:::||||||||
Db     3091 GCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCA 3150
```

```
Qy     387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Db    3151 GCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAA 3210

Qy     407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
Db    3211 GGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTAT 3270

Qy     427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db    3271 CCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAG 3330

Qy     447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
Db    3331 ACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAG 3390

Qy     467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db    3391 CGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTG 3450

Qy     487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db    3451 AATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGAC 3510

Qy     507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db    3511 TGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAG 3570

Qy     527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db    3571 ATTGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAA 3630

Qy     547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Db    3631 TATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATT 3690

Qy     567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Db    3691 GCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCA 3750

Qy     587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
Db    3751 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 3810

Qy     607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db    3811 CGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTG 3870

Qy     627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
Db    3871 GTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAA 3930

Qy     647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db    3931 GGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3963
```

RESULT 3
US-08-784-649A-5
; Sequence 5, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
;   APPLICANT: Sikic, Branimir I
;   APPLICANT: Chen, Gang
;   TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
;   TITLE OF INVENTION: CYCLOSPORIN MODULATION
;   NUMBER OF SEQUENCES: 5
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Fish & Richardson
;   STREET: 2200 Sand Hill Road
;   CITY: Menlo Park
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94025
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/784,649A
;   FILING DATE:
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Sherwood, Pamela J
;   REGISTRATION NUMBER: Reg.No. 5830697 36,677
;   REFERENCE/DOCKET NUMBER: 06037/007001
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415-322-5070
;   TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 5:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4264 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
US-08-784-649A-5

Alignment Scores:			
Pred. No.:	6.17e-219	Length:	4264
Score:	1872.00	Matches:	356
Percent Similarity:	74.22%	Conservative:	142
Best Local Similarity:	53.06%	Mismatches:	159
Query Match:	56.33%	Indels:	14
DB:	2	Gaps:	4

US-09-873-409-1 (1-659) x US-08-784-649A-5 (1-4264)

```
Qy       1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db    1951 GTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAA 2010

Qy      21 LeuValMetSerGln------------------AspIleLysLysAlaAspGluGlnMet 34
Db    2011 CTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCC 2070

Qy      35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
Db    2071 AAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGA 2130

Qy      54 ---------------LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
Db    2131 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACC 2190

Qy      68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db    2191 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAAT 2250

Qy      88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db    2251 TTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTG 2310

Qy     108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
Db    2311 CAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGAT 2370

Qy     127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db    2371 CCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATT 2430

Qy     147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db    2431 ATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTC 2490

Qy     167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
```

```
Db   2491 ACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 2550

Qy    187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
          |||:::   :::|||:::|||||   |||||||||  ||||||  |||   ||||||
Db   2551 GATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAA 2610

Qy    207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
          ::::::||||||   |||||||||:::   |||:::|||||||||      |||:::|||
Db   2611 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGG 2670

Qy    227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
             :::|||||||||||||||||||||||||:::::::|||   |||:::|||:::|||
Db   2671 ACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATT 2730

Qy    247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
                |||::::::|||:::   |||::::::|||          ::::::|||   |||
Db   2731 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 2790

Qy    267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
          |||||||||:::||||||:::   ||||||||||||||||||||||||:::||||||
Db   2791 AAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 2850

Qy    287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
          ||||||:::||||||||||||:::|||:::   ||||||   ||||||   :::   |||
Db   2851 CGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTG 2910

Qy    307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
          |||         :::||||||:::   :::||||||   |||   |||   ::::::|||
Db   2911 CAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTC 2970

Qy    327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
          :::   |||   :::||||||:::||||||         |||||||||||||||||||||:::
Db   2971 ACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTG 3030

Qy    347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
                      |||:::   |||   :::   :::||||||:::|||:::   :::|||
Db   3031 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTCAGCTGTTGTCTTTGGT 3090

Qy    367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
          |||||||||:::|||:::               ||||||:::|||:::|||||||||
Db   3091 GCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCA 3150

Qy    387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
          |||||||||:::      ::::::||||||   |||   |||||||||   |||   |||
Db   3151 GCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAA 3210

Qy    407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
          |||      |||:::|||   ||||||||||:::   |||   ||||||   |||   |||
Db   3211 GGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTAT 3270

Qy    427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
          |||   |||||||||:::   :::|||:::||||||||||||      :::::::::|||:::
Db   3271 CCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAG 3330

Qy    447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
          |||:::|||   ||||||||||||||||||||||||||||      ||||||||||||:::
Db   3331 ACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAG 3390

Qy    467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
          |||   |||||||||:::   |||:::||||||   ||||||   :::   |||   |||
Db   3391 CGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTG 3450

Qy    487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
          ||||||||||||||||||:::   :::   ||||||   |||||||||:::||||||:::
Db   3451 AATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGAC 3510

Qy    507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
          ||||||||||||||||||||||||||||||||||||||||||||||      :::|||
Db   3511 TGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAG 3570

Qy    527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
          |||      ||||||      ||||||||||||:::|||||||||   ||||||   |||
Db   3571 ATTGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAA 3630

Qy    547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
          |||:::|||:::||||||   ||||||   ||||||||||||||||||||||||:::
Db   3631 TATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATT 3690

Qy    567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
          |||||||||||||||||:::::::::|||   ||||||||||||||||||||||||
Db   3691 GCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCA 3750

Qy    587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
          |||||||||   :::||||||||||||||||||   ||||||||||||||||||   |||
Db   3751 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 3810

Qy    607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
          ||||||||:::|||:::   ||||||||||||   |||||||||||||||||||||
Db   3811 CGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTG 3870

Qy    627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
          |||      |||||:::::::||||||   ||||||||||||:::||||||      :::
Db   3871 GTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAA 3930

Qy    647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
             |||||||||   :::|||:::   |||:::
Db   3931 GGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3963
```

RESULT 4
US-08-181-471-2
; Sequence 2, Application US/08181471
; Patent No. 5641508
; GENERAL INFORMATION:
; APPLICANT: Li, Lingna
; APPLICANT: Lishko, Valeryi K.
; TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL
; TITLE OF INVENTION: COMPOSITIONS TO HAIR FOLLICLES
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thomas Fitting
; STREET: 12526 High Bluff Drive, Suite 300
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92130
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/181,471
; FILING DATE: 13-JAN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/041,553
; FILING DATE: 02-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: ANT0029P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-792-3680
; TELEFAX: 619-792-8477
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4646 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 425..4267

```
US-08-181-471-2

Alignment Scores:
Pred. No.:              7.18e-219    Length:         4646
Score:                  1872.00      Matches:        356
Percent Similarity:     74.22%       Conservative:   142
Best Local Similarity:  53.06%       Mismatches:     159
Query Match:            56.33%       Indels:         14
DB:                     1            Gaps:           4

US-09-873-409-1 (1-659) x US-08-181-471-2 (1-4646)

Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db   2237 GTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAA 2296

Qy     21 LeuValMetSerGln------------------AspIleLysLysAlaAspGluGlnMet 34
Db   2297 CTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCC 2356

Qy     35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
Db   2357 AAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGA 2416

Qy     54 ---------------LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
Db   2417 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACC 2476

Qy     68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db   2477 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAAT 2536

Qy     88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db   2537 TTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTG 2596

Qy    108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
Db   2597 CAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGAT 2656

Qy    127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db   2657 CCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATT 2716

Qy    147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db   2717 ATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTC 2776

Qy    167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
Db   2777 ACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 2836

Qy    187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db   2837 GATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAA 2896

Qy    207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db   2897 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGG 2956

Qy    227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
Db   2957 ACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATT 3016

Qy    247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db   3017 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 3076

Qy    267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db   3077 AAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 3136

Qy    287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db   3137 CGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTG 3196

Qy    307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db   3197 CAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTC 3256

Qy    327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
Db   3257 ACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTG 3316

Qy    347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db   3317 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGT 3376

Qy    367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db   3377 GCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCA 3436

Qy    387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Db   3437 GCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAA 3496

Qy    407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
Db   3497 GGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTAT 3556

Qy    427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db   3557 CCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAG 3616

Qy    447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
Db   3617 ACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAG 3676

Qy    467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db   3677 CGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTG 3736

Qy    487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db   3737 AATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGAC 3796

Qy    507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db   3797 TGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAG 3856

Qy    527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db   3857 ATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAA 3916

Qy    547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Db   3917 TATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATT 3976

Qy    567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Db   3977 GCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCA 4036

Qy    587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
Db   4037 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 4096

Qy    607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db   4097 CGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTG 4156

Qy    627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
Db   4157 GTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAA 4216

Qy    647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
```

```
Db    4217 GGCATCTATTTTCAATGGTCAGTGTCCAGGCT 4249

RESULT 5
5206352-3
;Patent No. 5206352
;    APPLICANT: Roninson, Igor B.;Pastan Ira H.;Gottesman,
;Michael M.
;    TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA
;SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
;    NUMBER OF SEQUENCES: 4
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/622,836
;      FILING DATE: 24-SEP-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 892,575
;      FILING DATE: 01-AUG-1986
;      APPLICATION NUMBER: 845,610
;      FILING DATE: 28-MAR-1986
;SEQ ID NO:3:
;      LENGTH: 4669
5206352-3

Alignment Scores:
Pred. No.:               7.24e-219     Length:         4669
Score:                   1872.00       Matches:        356
Percent Similarity:      74.22%        Conservative:   142
Best Local Similarity:   53.06%        Mismatches:     159
Query Match:             56.33%        Indels:         14
DB:                      6             Gaps:           4

US-09-873-409-1 (1-659) x 5206352-3 (1-4669)

Qy       1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db    2237 GTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAA 2296

Qy      21 LeuValMetSerGln-----------------AspIleLysLysAlaAspGluGlnMet 34
Db    2297 CTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCC 2356

Qy      35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
Db    2357 AAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGA 2416

Qy      54 ---------------LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
Db    2417 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACC 2476

Qy      68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db    2477 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAAT 2536

Qy      88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db    2537 TTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTG 2596

Qy     108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
Db    2597 CAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGAT 2656

Qy     127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db    2657 CCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATT 2716

Qy     147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db    2717 ATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTC 2776

Qy     167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
Db    2777 ACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 2836

Qy     187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db    2837 GATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAA 2896

Qy     207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db    2897 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGG 2956

Qy     227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
Db    2957 ACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATT 3016

Qy     247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db    3017 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 3076

Qy     267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db    3077 AAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 3136

Qy     287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db    3137 CGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTG 3196

Qy     307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db    3197 CAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTC 3256

Qy     327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
Db    3257 ACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTG 3316

Qy     347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db    3317 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGT 3376

Qy     367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db    3377 GCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCA 3436

Qy     387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Db    3437 GCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAA 3496

Qy     407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
Db    3497 GGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTAT 3556

Qy     427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db    3557 CCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAG 3616

Qy     447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
Db    3617 ACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAG 3676

Qy     467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db    3677 CGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTG 3736

Qy     487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db    3737 AATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGAC 3796

Qy     507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db    3797 TGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAG 3856

Qy     527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db    3857 ATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAA 3916

Qy     547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
```

```
Db    3917 TATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATT 3976

Qy     567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
           ||||||||||||||||:::::::::|||   ||||||||||||||||||||||||
Db    3977 GCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCA 4036

Qy     587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
           ||||||||   :::||||||||||||||||  ||||||||||||||||   |||
Db    4037 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 4096

Qy     607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
           ||||||||:::|||:::   ||||||||||   ||||||||||||||||||||||
Db    4097 CGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTG 4156

Qy     627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
           |||      |||||::::::|||||   ||||||||||:::||||||         :::
Db    4157 GTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAA 4216

Qy     647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
              ||||||||   :::|||:::   |||:::
Db    4217 GGCATCTATTTTCAATGGTCAGTGTCCAGGCT 4249
```

```
RESULT 6
US-08-793-610-5
; Sequence 5, Application US/08793610
; Patent No. 5858744
;  GENERAL INFORMATION:
;    APPLICANT:  BAUM, Christopher
;    APPLICANT:  STOCKING-HARBERS, Carol
;    APPLICANT:  OSTERTAG, Wolfram
;    TITLE OF INVENTION:  RETROVIRAL VECTOR HYBRIDS AND THE USE THEROF
;    TITLE OF INVENTION:  FOR GENE TRANSFER
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nikaido, Marmelstein, Murray & Oram LLP
;      STREET:  655 Fifteenth Street N.W. Suite 330
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/793,610
;      FILING DATE:  07-MAR-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P 44 31 973.8
;      FILING DATE:  08-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 195 03 952.1
;      FILING DATE:  07-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/03175
;      FILING DATE:  10-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berman, Richard J.
;      REGISTRATION NUMBER:  39,105
;      REFERENCE/DOCKET NUMBER:  P1614-7007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)638-5000
;      TELEFAX:  (202)638-4810
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6505 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA
US-08-793-610-5
```

```
Alignment Scores:
Pred. No.:                1.3e-218       Length:        6505
Score:                    1872.00        Matches:       356
Percent Similarity:       74.22%         Conservative:  142
Best Local Similarity:    53.06%         Mismatches:    159
Query Match:              56.33%         Indels:        14
DB:                       2              Gaps:          4

US-09-873-409-1 (1-659) x US-08-793-610-5 (1-6505)

Qy       1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
           ::::::   |||||||||   |||   |||||||||   :::::||||:::|||:::
Db    3629 GTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAA 3688

Qy      21 LeuValMetSerGln------------------AspIleLysLysAlaAspGluGlnMet 34
           ||||||   |||                     :::::::::   |||   ::::::
Db    3689 CTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCC 3748

Qy      35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
           :::|||         :::   |||   :::|||                     :::
Db    3749 AAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGA 3808

Qy      54 ---------------LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
                          :::|||::::::   |||         |||:::         :::
Db    3809 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACC 3868

Qy      68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
           ::::::   :::   |||:::|||   ||||||      :::|||:::||||||||||
Db    3869 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAAT 3928

Qy      88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
                 |||||||||:::   |||:::|||         :::::::::||||||   :::
Db    3929 TTAACTGAATGGCCTTATTTGTTGTTGGTGTATTTGTGCCATTATAAATGGAGGCCTG 3988

Qy     108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
              |||   |||:::|||||||||:::|||||||||   :::|||         :::|||
Db    3989 CAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTACAAGAATTGATGAT 4048

Qy     127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
                 |||   :::   ::::::   ::::::|||::::::|||:::   ||||||:::
Db    4049 CCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATT 4108

Qy     147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
           |||      |||:::::::::|||:::||||||      :::|||:::||||||||||||||
Db    4109 ATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTC 4168

Qy     167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
           |||   ||||||||||::::::   |||::::::||||||   ||||||::::::||||||
Db    4169 ACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 4228

Qy     187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
           |||:::   :::|||:::||||||   |||||||||   ||||||   |||   ||||||
Db    4229 GATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAA 4288

Qy     207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
           ::::::||||||   |||||||||:::   |||:::|||||||||      |||:::|||
Db    4289 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGG 4348

Qy     227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
              :::||||||||||||||||||||||||||||::::::|||   |||:::|||:::|||
Db    4349 ACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATT 4408

Qy     247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
              |||::::::|||:::   |||::::::|||         ::::::|||   |||
Db    4409 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 4468

Qy     267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
           |||||||||:::||||||:::   |||||||||||||||||||||||||:::||||||
Db    4469 AAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 4528

Qy     287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
           ||||||:::||||||||||||:::|||:::   ||||||   ||||||   :::   |||
Db    4529 CGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTG 4588
```

```
Qy    307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
          |||      :::||||||:::   :::||||||   |||   |||      ::::::|||
Db   4589 CAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTC 4648

Qy    327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
           :::   |||   :::||||||:::||||||      ||||||||||||||||||||:::
Db   4649 ACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTG 4708

Qy    347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
                      |||:::   |||   :::   :::||||||:::|||:::   :::|||
Db   4709 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTCAGCTGTTGTCTTTGGT 4768

Qy    367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
          ||||||||:::|||:::                  ||||||:::|||:::|||||||||
Db   4769 GCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCA 4828

Qy    387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
          |||||||||:::         ::::::||||||   |||   |||||||||   |||   |||
Db   4829 GCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAA 4888

Qy    407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
          |||      |||:::|||   |||||||||:::   |||   ||||||   |||   |||
Db   4889 GGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTAT 4948

Qy    427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
          |||   |||||||||:::   :::|||:::|||||||||||||   :::::::::|||:::
Db   4949 CCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAG 5008

Qy    447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
          |||:::|||   ||||||||||||||||||||||||||||||   ||||||||||||:::
Db   5009 ACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAG 5068

Qy    467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
          |||   |||||||||:::   |||:::||||||         :::   |||   |||
Db   5069 CGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTG 5128

Qy    487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
          ||||||||||||||||||:::   :::   ||||||   |||||||||:::||||||:::
Db   5129 AATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGAC 5188

Qy    507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
          ||||||||||||||||||||||||||||||||||||||||||||||||      :::|||
Db   5189 TGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAG 5248

Qy    527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
          |||         ||||||      ||||||||||||:::|||||||||   ||||||   |||
Db   5249 ATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAA 5308

Qy    547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
          |||:::|||:::|||      ||||||   |||||||||||||||||||||||||||:::
Db   5309 TATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATT 5368

Qy    567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
          ||||||||||||||||||:::::::::|||   ||||||||||||||||||||||||||
Db   5369 GCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCA 5428

Qy    587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
          ||||||||   :::|||||||||||||||||||   ||||||||||||||||||   |||
Db   5429 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 5488

Qy    607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
          |||||||||:::|||:::   ||||||||||||   ||||||||||||||||||||||||
Db   5489 CGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTG 5548

Qy    627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
          |||      ||||||:::::::||||||   ||||||||||||:::||||||      :::
Db   5549 GTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAA 5608

Qy    647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
          |||||||||   :::|||:::   |||:::
Db   5609 GGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 5641
```

```
RESULT 7
US-08-793-610-6
; Sequence 6, Application US/08793610
; Patent No. 5858744
;  GENERAL INFORMATION:
;    APPLICANT:  BAUM, Christopher
;    APPLICANT:  STOCKING-HARBERS, Carol
;    APPLICANT:  OSTERTAG, Wolfram
;    TITLE OF INVENTION:  RETROVIRAL VECTOR HYBRIDS AND THE USE THEROF
;    TITLE OF INVENTION:  FOR GENE TRANSFER
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nikaido, Marmelstein, Murray & Oram LLP
;      STREET:  655 Fifteenth Street N.W. Suite 330
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/793,610
;      FILING DATE:  07-MAR-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P 44 31 973.8
;      FILING DATE:  08-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 195 03 952.1
;      FILING DATE:  07-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/03175
;      FILING DATE:  10-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berman, Richard J.
;      REGISTRATION NUMBER:  39,105
;      REFERENCE/DOCKET NUMBER:  P1614-7007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)638-5000
;      TELEFAX:  (202)638-4810
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9318 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA
US-08-793-610-6

Alignment Scores:
Pred. No.:               2.45e-218      Length:         9318
Score:                   1872.00        Matches:        356
Percent Similarity:      74.22%         Conservative:   142
Best Local Similarity:   53.06%         Mismatches:     159
Query Match:             56.33%         Indels:         14
DB:                      2              Gaps:           4

US-09-873-409-1 (1-659) x US-08-793-610-6 (1-9318)

Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
             ::::::   |||||||||   |||   |||||||||   ::::::|||:::|||:::
Db   3588 GTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAA 3647

Qy     21 LeuValMetSerGln------------------AspIleLysLysAlaAspGluGlnMet 34
             ||||||   |||                      :::::::::   |||   ::::::
Db   3648 CTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCC 3707

Qy     35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
             :::|||      :::   |||      :::|||                     :::
Db   3708 AAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGA 3767
```

```
Qy     54 ---------------LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
Db   3768 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACC 3827

Qy     68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
Db   3828 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAAT 3887

Qy     88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
Db   3888 TTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTG 3947

Qy    108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
Db   3948 CAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGAT 4007

Qy    127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
Db   4008 CCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATT 4067

Qy    147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
Db   4068 ATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTC 4127

Qy    167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
Db   4128 ACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 4187

Qy    187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
Db   4188 GATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAA 4247

Qy    207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
Db   4248 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGG 4307

Qy    227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
Db   4308 ACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATT 4367

Qy    247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
Db   4368 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 4427

Qy    267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db   4428 AAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 4487

Qy    287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db   4488 CGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTG 4547

Qy    307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db   4548 CAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTC 4607

Qy    327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
Db   4608 ACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTG 4667

Qy    347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db   4668 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGT 4727

Qy    367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db   4728 GCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCA 4787

Qy    387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Db   4788 GCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAA 4847

Qy    407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
Db   4848 GGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTAT 4907

Qy    427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db   4908 CCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAG 4967

Qy    447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
Db   4968 ACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAG 5027

Qy    467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db   5028 CGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTG 5087

Qy    487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db   5088 AATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGAC 5147

Qy    507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db   5148 TGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAG 5207

Qy    527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db   5208 ATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAA 5267

Qy    547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Db   5268 TATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATT 5327

Qy    567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Db   5328 GCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCA 5387

Qy    587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
Db   5388 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 5447

Qy    607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db   5448 CGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTG 5507

Qy    627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
Db   5508 GTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAA 5567

Qy    647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db   5568 GGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 5600
```

```
RESULT 8
US-08-752-447-1
; Sequence 1, Application US/08752447
; Patent No. 5994088
;  GENERAL INFORMATION:
;    APPLICANT:  Mechetner, Eugene
;    APPLICANT:  Roninson, Igor B
;    TITLE OF INVENTION:  Methods and Reagents for Preparing and
;    TITLE OF INVENTION:  Using Immunoligcal Agents Specific for P-glycoprotein
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McDonnell Boehnen Hulbert & Berghoff Ltd.
;      STREET:  300 South Wacker Drive, Seventh Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/752,447
;      FILING DATE:  15-NOV-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5994088nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  95,1121
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-913-0001
;      TELEFAX:  312-913-9808
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4669 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  5'UTR
;      LOCATION:  1..424
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  425..4264
;    FEATURE:
;      NAME/KEY:  3'UTR
;      LOCATION:  4265..4669
US-08-752-447-1

Alignment Scores:
Pred. No.:              8.88e-217     Length:         4669
Score:                  1855.00       Matches:        354
Percent Similarity:     73.47%        Conservative:   139
Best Local Similarity:  52.76%        Mismatches:     164
Query Match:            55.82%        Indels:         14
DB:                     2             Gaps:           4

US-09-873-409-1 (1-659) x US-08-752-447-1 (1-4669)

Qy       1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
             ::::::    ||||||||   |||   |||||||||    ::::::||::||||:::
Db    2237 GTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAA 2296

Qy      21 LeuValMetSerGln-----------------AspIleLysLysAlaAspGluGlnMet 34
           ||||||   |||                    ::::::::   |||     ::::::
Db    2297 CTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCC 2356

Qy      35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
           :::|||         :::   |||      :::|||                  :::
Db    2357 AAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGA 2416

Qy      54 ---------------LysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr 68
                          :::|||::::::                  :::    ||||||
Db    2417 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAACACAGAAAGCTTAGTACC 2476

Qy      69 ---GlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
              ::::::   :::   |||:::|||   ||||||        :::|||:::|||||||||
Db    2477 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAAT 2536

Qy      88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
                    ||||||||:::   |||:::|||         :::::::::||||||    :::
Db    2537 TTAACTGAATGGCCTTATTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTG 2596

Qy     108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
              |||   |||:::|||||||||:::|||||||||      :::|||          :::|||
Db    2597 CAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTACAAGAATTGATGAT 2656

Qy     127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
                    |||   :::   ::::::   ::::::|||::::::|||:::   ||||||:::
Db    2657 CCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATT 2716

Qy     147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
           |||   |||:::::::::|||:::||||||         :::|||:::|||||||||||||||
Db    2717 ATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTC 2776

Qy     167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
           |||   |||||||||::::::   |||::::::||||||   ||||||::::::||||||
Db    2777 ACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 2836

Qy     187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
              :::   :::|||:::||||||   |||||||||   ||||||   |||   ||||||
Db    2837 CATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAA 2896

Qy     207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
           ::::::||||||   |||||||||:::   |||:::|||||||||         |||:::|||
Db    2897 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGG 2956

Qy     227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
              :::|||||||||||||||||||||||||||::::::|||   |||:::|||:::|||
Db    2957 ACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATT 3016

Qy     247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
              |||::::::|||:::   |||::::::|||                  |||   |||
Db    3017 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTTGCTGGACAAGCACTG 3076

Qy     267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
           |||||||||:::||||||:::   |||||||||||||||||||||||||:::||||||
Db    3077 AAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 3136

Qy     287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
           ||||||:::||||||||||||:::|||:::      ||||||   ||||||   :::   |||
Db    3137 CGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTG 3196

Qy     307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
           |||      :::||||||:::   :::||||||   |||   |||      ::::::|||
Db    3197 CAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTC 3256

Qy     327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
           :::   |||   :::||||||:::||||||      ||||||||||||||||||||:::
Db    3257 ACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTG 3316

Qy     347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
                       |||:::   |||   :::   :::||||||:::|||:::   :::|||
Db    3317 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGT 3376

Qy     367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
           |||||||||:::|||:::               ||||||:::|||:::|||||||||
Db    3377 GCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCA 3436

Qy     387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
           |||||||||:::         ::::::||||||   |||   |||||||||   |||   |||
Db    3437 GCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAA 3496

Qy     407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
           |||         |||:::|||   |||||||||:::   |||   ||||||   |||   |||
Db    3497 GGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTAT 3556

Qy     427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
           |||   |||||||||:::   :::|||:::||||||||||||   :::::::::|||:::
Db    3557 CCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAG 3616

Qy     447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
           |||:::|||   ||||||||||||||||||||||||||||||   ||||||||||||:::
Db    3617 ACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAG 3676

Qy     467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
           |||   ||||||||||:::   |||:::||||||   ||||||   :::   |||   |||
Db    3677 CGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTG 3736

Qy     487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
           ||||||||||||||||:::   :::   ||||||   |||||||||:::||||||:::
Db    3737 AATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGAC 3796

Qy     507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
           |||||||||||||||||||||||||||||||||||||||||||||||      :::|||
Db    3797 TGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAG 3856
```

```
Qy     527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
               |||      ||||||     |||||||||||:::|||||||||   ||||||    |||
Db    3857 ATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAA 3916

Qy     547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
               |||:::|||:::||||||   ||||||   ||||||||||||||||||||||||:::
Db    3917 TATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATT 3976

Qy     567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
               ||||||||||||   |||:::::::::|||   ||||||||||||||||||||||||
Db    3977 GCCATAGCTCGTCGCCTTGTTAGACAGCCTCATATTTGCTTTTGGATGAAGCCACGTCA 4036

Qy     587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
               |||||||||   :::||||||||||||||||||   ||||||||||||||||   |||
Db    4037 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 4096

Qy     607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
               |||||||||:::|||:::   ||||||||||||   ||||||||||||||||||||||
Db    4097 CGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTG 4156

Qy     627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
               |||      ||||||::::::||||||   ||||||||||||:::||||||      :::
Db    4157 GTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAA 4216

Qy     647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
                  |||||||||   :::|||:::   |||:::
Db    4217 GGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 4249

RESULT 9
US-09-316-167-1
; Sequence 1, Application US/09316167
; Patent No. 6365357
;  GENERAL INFORMATION:
;    APPLICANT:  Mechetner, Eugene
;    APPLICANT:  Roninson, Igor B
;    TITLE OF INVENTION:  Methods and Reagents for Preparing and
;    TITLE OF INVENTION:  Using Immunoligcal Agents Specific for P-glycoprotein
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McDonnell Boehnen Hulbert & Berghoff Ltd.
;      STREET:  300 South Wacker Drive, Seventh Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/316,167
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/752,447
;      FILING DATE:  15-NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6365357nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  95,1121
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-913-0001
;      TELEFAX:  312-913-9808
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4669 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  5'UTR
;      LOCATION:  1..424
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  425..4264
;    FEATURE:
;      NAME/KEY:  3'UTR
;      LOCATION:  4265..4669
US-09-316-167-1

Alignment Scores:
Pred. No.:              8.88e-217      Length:        4669
Score:                  1855.00        Matches:       354
Percent Similarity:     73.47%         Conservative:  139
Best Local Similarity:  52.76%         Mismatches:    164
Query Match:            55.82%         Indels:        14
DB:                     4              Gaps:          4

US-09-873-409-1 (1-659) x US-09-316-167-1 (1-4669)

Qy       1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
           ::::::   |||||||||   |||   |||||||||   ::::::|||:::|||:::
Db    2237 GTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAA 2296

Qy      21 LeuValMetSerGln------------------AspIleLysLysAlaAspGluGlnMet 34
           ||||||   |||                     :::::::::   |||   ::::::
Db    2297 CTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCC 2356

Qy      35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
           :::|||      :::   |||      :::|||                     :::
Db    2357 AAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGA 2416

Qy      54 ---------------LysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr 68
                          :::|||::::::                     :::   ||||||
Db    2417 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAACACAGAAAGCTTAGTACC 2476

Qy      69 ---GlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
              ::::::   :::   |||:::|||   ||||||      :::|||:::|||||||||
Db    2477 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAAT 2536

Qy      88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
                 |||||||||:::   |||:::|||      :::::::::||||||   :::
Db    2537 TTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTG 2596

Qy     108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
              |||   |||:::|||||||||:::|||||||||      :::|||        :::|||
Db    2597 CAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGAT 2656

Qy     127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
                 |||   :::   ::::::   ::::::|||::::::|||:::   ||||||:::
Db    2657 CCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATT 2716

Qy     147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
           |||   |||:::::::::|||:::||||||      :::|||:::|||||||||||||||
Db    2717 ATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTC 2776

Qy     167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
           |||   ||||||||||::::::   |||::::::||||||   |||||::::::||||||
Db    2777 ACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 2836

Qy     187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
              :::   :::|||:::||||||   |||||||||   ||||||   |||   ||||||
Db    2837 CATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAA 2896

Qy     207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
           ::::::||||||   |||||||||:::   |||:::|||||||||      |||:::|||
Db    2897 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGG 2956

Qy     227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
              :::||||||||||||||||||||||||||||:::::::|||   |||:::|||:::|||
Db    2957 ACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATT 3016

Qy     247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
```

```
Db   3017 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTTGCTGGACAAGCACTG 3076

Qy    267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
Db   3077 AAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 3136

Qy    287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
Db   3137 CGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTG 3196

Qy    307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
Db   3197 CAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTC 3256

Qy    327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
Db   3257 ACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTG 3316

Qy    347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
Db   3317 GCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGT 3376

Qy    367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
Db   3377 GCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCA 3436

Qy    387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
Db   3437 GCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAA 3496

Qy    407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
Db   3497 GGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTAT 3556

Qy    427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
Db   3557 CCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAG 3616

Qy    447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
Db   3617 ACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAG 3676

Qy    467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
Db   3677 CGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTG 3736

Qy    487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
Db   3737 AATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGAC 3796

Qy    507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
Db   3797 TGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAG 3856

Qy    527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
Db   3857 ATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAA 3916

Qy    547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
Db   3917 TATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATT 3976

Qy    567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
Db   3977 GCCATAGCTCGTCGCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCA 4036

Qy    587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
Db   4037 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 4096

Qy    607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
Db   4097 CGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTG 4156

Qy    627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
Db   4157 GTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAA 4216

Qy    647 AspIleTyrPheLysLeuValAsnAlaGlnSer 657
Db   4217 GGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 4249
```

RESULT 10
US-09-120-513-1
; Sequence 1, Application US/09120513
; Patent No. 6025160
; GENERAL INFORMATION:
; APPLICANT: Brun, Kimberly
; APPLICANT: Chenery, Richard
; APPLICANT: Ellens, Harma
; APPLICANT: Field, John
; APPLICANT: Yue, Lin
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE
; TITLE OF INVENTION: SEQUENCES ENCODING RAT MDR1B2 AND
; TITLE OF INVENTION: SCREENING METHODS THEREOF
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY:
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/120,513
; FILING DATE: 22-JUL-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: King, William T
; REGISTRATION NUMBER: 30,954
; REFERENCE/DOCKET NUMBER: GP50008
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5015
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4233 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-09-120-513-1

Alignment Scores:

Pred. No.:	3.54e-216	Length:	4233
Score:	1849.50	Matches:	358
Percent Similarity:	73.13%	Conservative:	132
Best Local Similarity:	53.43%	Mismatches:	163
Query Match:	55.66%	Indels:	17
DB:	3	Gaps:	5

US-09-873-409-1 (1-659) x US-09-120-513-1 (1-4233)

```
Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db   1832 GTCATTGTGGAGCAAGGAAATCATGAAGAGCTCATGAAAGAGAAGGGCATTTACTTCAAA 1891

Qy     21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
```

```
Db   1892 CTTGTCATGACACAG---ACTAGAGGAAATGAAATTGAACCAGGAAATAATGCTTATGAA 1948

Qy     41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
Db   1949 TCCCAAAGTGACACTGGTGCCTCTGAGTTGACTTCAGAAAAATCA---AAATCTCCTTTA 2005

Qy     61 IleAspLysAla--------------------------------GluGluSerThrGln 69
Db   2006 ATAAGGAGATCAATTCGCAGAAGTATCCACAGAAGACAAGACCAGGAGAGAAGACTTAGT 2065

Qy     70 SerLysGlu---------IleSerLeuProGluValSerLeuLeuLysIleLeuLysLeu 86
Db   2066 TCGAAAGAGGATGTGGATGAAGATGTGCCTATGGTTTCCTTTTGGCAGATCCTAAAGCTA 2125

Qy     87 AsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThr 106
Db   2126 AATATTAGTGAATGGCCCTATTTAGTTGTGGGTGTACTTTGTGCTGTTATAAATGGGTGC 2185

Qy    107 ValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGly---AsnAsn 125
Db   2186 ATACAACCAGTGTTTGCCATAGTGTTTTCAAAGATTGTAGGGGTTTTTCAAGAGACGAC 2245

Qy    126 AspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGly 145
Db   2246 GACCATGAAACCAAACAACGGAATTGTAACTTGTTTTCCCTTCTCTTTCTGGTCATGGGA 2305

Qy    146 ValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIle 165
Db   2306 ATGATTTCTTTTGTTACGTACTTCTTTCAAGGCTTCACATTTGGCAAAGCTGGAGAGATC 2365

Qy    166 LeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrp 185
Db   2366 CTCACCAAGCGACTCCGATACATGGTCTTCAAATCCATGCTGCGACAGGATATAAGCTGG 2425

Qy    186 PheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAla 205
Db   2426 TTTGATGACCATAAAAACACCACTGGCTCGCTGACTACCAGGCTCGCTAGTGACGCTTCT 2485

Qy    206 GlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMet 225
Db   2486 AATGTTAAAGGGGCTATGGGCTCCAGGCTTGCTGTAGTTACCCAGAATGTAGCAAACCTT 2545

Qy    226 GlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSer 245
Db   2546 GGCACAGGAATTATCTTATCCTTAGTCTATGGCTGGCAGCTTACACTTTTACTTGTAGTA 2605

Qy    246 IleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAla 265
Db   2606 ATTATACCACTCATTGTCTTGGGTGGAATTATTGAAATGAAACTGTTGTCTGGTCAAGCC 2665

Qy    266 AsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsn 285
Db   2666 TTGAAGGACAAGAAAGAGCTAGAGATCTCTGGGAAGATCGCTACAGAAGCAATTGAAAAC 2725

Qy    286 IleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMet 305
Db   2726 TTCCGCACTGTTGTCTCTTTGACTAGGGAGCAGAAGTTTGAAACTATGTATGCCCAGAGC 2785

Qy    306 LeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAla 325
Db   2786 TTGCAGATACCATACAGAAATGCTTTGAAGAAAGCACACGTCTTTGGGATCACCTTCGCC 2845

Qy    326 PheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeu 345
Db   2846 TTCACCCAGGCCATGATTTATTTTTCCTATGCTGCTTGTTTCCGGTTCGGTGCCTACTTG 2905

Qy    346 IleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyr 365
Db   2906 GTGGCACGAGAACTCATGACGTTTGAAAATGTTATGTTGGTATTTTCTGCTGTTGTCTTT 2965

Qy    366 GlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSer 385
Db   2966 GGTGCCATGGCAGCAGGGAATACCAGTTCATTCGCTCCTGACTACGCGAAGGCCAAAGTC 3025

Qy    386 GlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGln 405
Db   3026 TCAGCATCCCACATCATCAGGATCATTGAGAAAATCCCCGAGATTGACAGCTACAGCACG 3085

Qy    406 GluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePhe 425
Db   3086 GAGGGCTTGAAGCCTAATTGGTTAGAAGGAAATGTGAAATTTAATGGAGTCATGTTCAAC 3145

Qy    426 TyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGly 445
Db   3146 TATCCCACCCGACCCAACATCCCAGTGCTTCAGGGACTGAGCTTCGAGGTGAAGAAGGGG 3205

Qy    446 LysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeu 465
Db   3206 CAAACGCTTCGCCTGGTGGGCAGCAGTGGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTC 3265

Qy    466 GlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGlu 485
Db   3266 GAGCGCTTCTACAACCCCATGGCTGGAACAGTGTTTCTAGATGGCAAAGAAATAAAACAA 3325

Qy    486 LeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPhe 505
Db   3326 CTCAACGTCCAGTGGCTCCGCGCCCACCTGGGCATTGTGTCCCAGGAGCCCATCCTGTTT 3385

Qy    506 AsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAsp 525
Db   3386 GACTGCAGCATCACCGAGAACATCGCCTACGGAGACAACAGCCGTGTCGTGTCTCATGAG 3445

Qy    526 GluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGlu 545
Db   3446 GAGATCGTGAGGGCCGCCAGGGAGGCCAACATCCACCAGTTCATCGACTCACTGCCTGAG 3505

Qy    546 LysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArg 565
Db   3506 AAATACAACACCAGAGTGGGAGACAAAGGGACTCAGCTGTCGGGCGGGCAGAAGCAGCGC 3565

Qy    566 LeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThr 585
Db   3566 ATCGCCATCGCGCGCGCCCTCGTCAGACAGCCTCACATCTTACTTCTGGATGAAGCGACA 3625

Qy    586 SerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThr 605
Db   3626 TCAGCTCTGGATACGGAGAGTGAAAAGGTCGTCCAGGAAGCGCTGGACAAAGCCAGGGAA 3685

Qy    606 GlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIle 625
Db   3686 GGCCGCACCTGCATTGTGATCGCGCACCGCCTGTCCACCATCCAGAACGCAGACTTGATC 3745

Qy    626 ValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsn 645
Db   3746 GTGGTGATTCAGAACGGCCAGGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCCCAG 3805

Qy    646 ArgAspIleTyrPheLysLeuValAsnAla 655
Db   3806 AAAGGCATCTATTTCTCGATGGTTCAGGCT 3835

RESULT 11
US-09-450-105-1
; Sequence 1, Application US/09450105
; Patent No. 6169166
; GENERAL INFORMATION:
;  APPLICANT: Kimberly Anne Brun
;  APPLICANT: Richard James Chenery
;  APPLICANT: Harma Ellens
;  APPLICANT: John Anthony Feild
;  APPLICANT: Lin Yue
;  TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
;  TITLE OF INVENTION: ENCODING RAT MDR1B2 AND SCREENING METHODS THEREOF
;  FILE REFERENCE: GP-50008-D1
;  CURRENT APPLICATION NUMBER: US/09/450,105
;  CURRENT FILING DATE: 1999-11-29
;  EARLIER APPLICATION NUMBER: 09/120,513
```

```
;  EARLIER FILING DATE: 1998-07-22
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 4233
;   TYPE: DNA
;   ORGANISM: HOMO SAPIENS
US-09-450-105-1

Alignment Scores:
Pred. No.:              3.54e-216     Length:        4233
Score:                  1849.50       Matches:       358
Percent Similarity:     73.13%        Conservative:  132
Best Local Similarity:  53.43%        Mismatches:    163
Query Match:            55.66%        Indels:        17
DB:                     4             Gaps:          5

US-09-873-409-1 (1-659) x US-09-450-105-1 (1-4233)

Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
Db   1832 GTCATTGTGGAGCAAGGAAATCATGAAGAGCTCATGAAAGAGAAGGGCATTTACTTCAAA 1891

Qy     21 LeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSer 40
Db   1892 CTTGTCATGACACAG---ACTAGAGGAAATGAAATTGAACCAGGAAATAATGCTTATGAA 1948

Qy     41 ThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe 60
Db   1949 TCCCAAAGTGACACTGGTGCCTCTGAGTTGACTTCAGAAAAATCA---AAATCTCCTTTA 2005

Qy     61 IleAspLysAla--------------------------------GluGluSerThrGln 69
Db   2006 ATAAGGAGATCAATTCGCAGAAGTATCCACAGAAGACAAGACCAGGAGAGAAGACTTAGT 2065

Qy     70 SerLysGlu---------IleSerLeuProGluValSerLeuLeuLysIleLeuLysLeu 86
Db   2066 TCGAAAGAGGATGTGGATGAAGATGTGCCTATGGTTTCCTTTTGGCAGATCCTAAAGCTA 2125

Qy     87 AsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThr 106
Db   2126 AATATTAGTGAATGGCCCTATTTAGTTGTGGGTGTACTTTGTGCTGTTATAAATGGGTGC 2185

Qy    107 ValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGly---AsnAsn 125
Db   2186 ATACAACCAGTGTTTGCCATAGTGTTTTCAAAGATTGTAGGGGTTTTTTCAAGAGACGAC 2245

Qy    126 AspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGly 145
Db   2246 GACCATGAAACCAAACAACGGAATTGTAACTTGTTTTCCCTTCTCTTTCTGGTCATGGGA 2305

Qy    146 ValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIle 165
Db   2306 ATGATTTCTTTTGTTACGTACTTCTTTCAAGGCTTCACATTTGGCAAAGCTGGAGAGATC 2365

Qy    166 LeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrp 185
Db   2366 CTCACCAAGCGACTCCGATACATGGTCTTCAAATCCATGCTGCGACAGGATATAAGCTGG 2425

Qy    186 PheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAla 205
Db   2426 TTTGATGACCATAAAAACACCACTGGCTCGCTGACTACCAGGCTCGCTAGTGACGCTTCT 2485

Qy    206 GlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMet 225
Db   2486 AATGTTAAAGGGGCTATGGGCTCCAGGCTTGCTGTAGTTACCCAGAATGTAGCAAACCTT 2545

Qy    226 GlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSer 245
Db   2546 GGCACAGGAATTATCTTATCCTTAGTCTATGGCTGGCAGCTTACACTTTTACTTGTAGTA 2605

Qy    246 IleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAla 265
Db   2606 ATTATACCACTCATTGTCTTGGGTGGAATTATTGAAATGAAACTGTTGTCTGGTCAAGCC 2665

Qy    266 AsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsn 285
Db   2666 TTGAAGGACAAGAAAGAGCTAGAGATCTCTGGGAAGATCGCTACAGAAGCAATTGAAAAC 2725

Qy    286 IleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMet 305
Db   2726 TTCCGCACTGTTGTCTCTTTGACTAGGGAGCAGAAGTTTGAAACTATGTATGCCCAGAGC 2785

Qy    306 LeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAla 325
Db   2786 TTGCAGATACCATACAGAAATGCTTTGAAGAAAGCACACGTCTTTGGGATCACCTTCGCC 2845

Qy    326 PheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeu 345
Db   2846 TTCACCCAGGCCATGATTTATTTTTCCTATGCTGCTTGTTTCCGGTTCGGTGCCTACTTG 2905

Qy    346 IleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyr 365
Db   2906 GTGGCACGAGAACTCATGACGTTTGAAAATGTTATGTTGGTATTTTCTGCTGTTGTCTTT 2965

Qy    366 GlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSer 385
Db   2966 GGTGCCATGGCAGCAGGGAATACCAGTTCATTCGCTCCTGACTACGCGAAGGCCAAAGTC 3025

Qy    386 GlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGln 405
Db   3026 TCAGCATCCCACATCATCAGGATCATTGAGAAAATCCCCGAGATTGACAGCTACAGCACG 3085

Qy    406 GluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePhe 425
Db   3086 GAGGGCTTGAAGCCTAATTGGTTAGAAGGAAATGTGAAATTTAATGGAGTCATGTTCAAC 3145

Qy    426 TyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGly 445
Db   3146 TATCCCACCCGACCCAACATCCCAGTGCTTCAGGGACTGAGCTTCGAGGTGAAGAAGGGG 3205

Qy    446 LysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeu 465
Db   3206 CAAACGCTTCGCCTGGTGGGCAGCAGTGGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTC 3265

Qy    466 GlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGlu 485
Db   3266 GAGCGCTTCTACAACCCCATGGCTGGAACAGTGTTTCTAGATGGCAAAGAAATAAAACAA 3325

Qy    486 LeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPhe 505
Db   3326 CTCAACGTCCAGTGGCTCCGCGCCCACCTGGGCATTGTGTCCCAGGAGCCCATCCTGTTT 3385

Qy    506 AsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAsp 525
Db   3386 GACTGCAGCATCACCGAGAACATCGCCTACGGAGACAACAGCCGTGTCGTGTCTCATGAG 3445

Qy    526 GluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGlu 545
Db   3446 GAGATCGTGAGGGCCGCCAGGGAGGCCAACATCCACCAGTTCATCGACTCACTGCCTGAG 3505

Qy    546 LysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArg 565
Db   3506 AAATACAACACCAGAGTGGGAGACAAAGGGACTCAGCTGTCGGGCGGGCAGAAGCAGCGC 3565

Qy    566 LeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThr 585
Db   3566 ATCGCCATCGCGCGCGCCCTCGTCAGACAGCCTCACATCTTACTTCTGGATGAAGCGACA 3625

Qy    586 SerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThr 605
Db   3626 TCAGCTCTGGATACGGAGAGTGAAAAGGTCGTCCAGGAAGCGCTGGACAAAGCCAGGGAA 3685

Qy    606 GlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIle 625
Db   3686 GGCCGCACCTGCATTGTGATCGCGCACCGCCTGTCCACCATCCAGAACGCAGACTTGATC 3745
```

```
Qy    626 ValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsn 645
          ||||||:::  ||||||::::::|||||  |||||||||||:::||||||
Db   3746 GTGGTGATTCAGAACGGCCAGGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCCCAG 3805

Qy    646 ArgAspIleTyrPheLysLeuValAsnAla 655
          :::   |||||||||   :::|||   |||
Db   3806 AAAGGCATCTATTTCTCGATGGTTCAGGCT 3835
```

RESULT 12
US-08-583-276-18
; Sequence 18, Application US/08583276
; Patent No. 5837536
; GENERAL INFORMATION:
;   APPLICANT: McDonagh, Kevin T.
;   APPLICANT: Nienhuis, Arthur
;   APPLICANT: Tolstoshev, Paul
;   TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN
;   TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED
;   TITLE OF INVENTION: SELECTION OF CELLS TRANSDUCED WITH SUCH GENES
;   NUMBER OF SEQUENCES: 19
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
;   ADDRESSEE: Cecchi & Stewart
;   STREET: 6 Becker Farm Road
;   CITY: Roseland
;   STATE: New Jersey
;   COUNTRY: USA
;   ZIP: 07068
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3.5 inch diskette
;   COMPUTER: IBM PS/2
;   OPERATING SYSTEM: PC-DOS
;   SOFTWARE: DW4.V2
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/583,276
;   FILING DATE: 05-JAN-1996
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/332,444
;   FILING DATE: 31-OCT-1994
;   APPLICATION NUMBER: 07/887,712
;   FILING DATE: 22-MAY-1992
; INFORMATION FOR SEQ ID NO: 18:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4669 bases
;   TYPE: nucleic acid
;   STRANDEDNESS: singular
;   TOPOLOGY: linear
;   MOLECULE TYPE:
;   DESCRIPTION: Genomic DNA
US-08-583-276-18

```
Alignment Scores:
Pred. No.:              1.5e-215      Length:         4669
Score:                  1845.00       Matches:        352
Percent Similarity:     73.58%        Conservative:   141
Best Local Similarity:  52.54%        Mismatches:     163
Query Match:            55.52%        Indels:         14
DB:                     2             Gaps:           4
```

US-09-873-409-1 (1-659) x US-08-583-276-18 (1-4669)

```
Qy      1 MetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSer 20
          ::::::   |||||||||   |||   |||||||||   ::::::|||:::|||:::
Db   2237 GTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAA 2296

Qy     21 LeuValMetSerGln------------------AspIleLysLysAlaAspGluGlnMet 34
          ||||||      |||                  :::::::::   |||   ::::::
Db   2297 CTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCC 2356

Qy     35 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--- 53
          :::|||      :::   |||      :::|||                     :::
Db   2357 AAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCACGTCTAATAAGA 2416

Qy     54 ---------------LysSerIleLys---SerAspPheIleAspLysAlaGluGluSer 67
                         :::|||::::::   |||      |||:::         :::
Db   2417 AAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAACCCCAAGACAGAAAGCTTAGTACC 2476

Qy     68 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 87
             ::::::   :::   |||:::|||   ||||||      :::|||:::|||||||||
Db   2477 AAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAAT 2536

Qy     88 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 107
                |||||||||:::   |||:::|||      :::::::::||||||   :::
Db   2537 TTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTGTGCCATTATAAATGGAGGCCTG 2596

Qy    108 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 126
             |||   |||:::||||||||||:::|||||||||      :::|||      :::|||
Db   2597 CAACCAGCATTTGCAATAATATTTCAAAGATTATAGGGGTTTTACAAGAATTGATGAT 2656

Qy    127 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 146
                |||   :::   :::      ::::::|||::::::|||:::   ||||||:::
Db   2657 CCTGAAACAAAACGACAGAATACTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATT 2716

Qy    147 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 166
          |||   |||:::::::::|||:::||||||      :::|||:::||||||||||||||
Db   2717 ATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTC 2776

Qy    167 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 186
          |||   |||||||||::::::      ||||||::::::||||||
Db   2777 ACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTT 2836

Qy    187 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 206
          |||:::   :::|||:::||||||   |||||||||   ||||||   |||   ||||||
Db   2837 GATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAA 2896

Qy    207 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 226
          ::::::||||||   |||||||||:::   |||:::|||||||||         :::|||
Db   2897 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAATTCTTGGG 2956

Qy    227 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 246
             :::||||||||||||||||||||||||||::::::|||   |||:::|||:::|||
Db   2957 ACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATT 3016

Qy    247 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 266
             |||::::::|||:::   |||::::::|||      ::::::|||   |||
Db   3017 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 3076

Qy    267 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 286
          |||||||||:::||||||:::   ||||||||||||||||||||||:::||||||
Db   3077 AAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTC 3136

Qy    287 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 306
          ||||||:::||||||||||||:::|||:::   ||||||   ||||||   :::   |||
Db   3137 CGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTG 3196

Qy    307 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 326
          |||      :::||||||:::   :::||||||   |||   |||      ::::::|||
Db   3197 CAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTC 3256

Qy    327 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 346
          :::   |||   :::||||||:::|||      |||||||||||||||||||:::
Db   3257 ACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTG 3316

Qy    347 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 366
                   |||:::   |||   :::   :::||||||:::|||:::   :::|||
Db   3317 GCACATAAACTAATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGT 3376

Qy    367 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 386
          |||||||||:::|||:::               ||||||:::|||:::|||||||||
Db   3377 GCCATGGCCGTGGGGCAAGTGAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCA 3436

Qy    387 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 406
          |||||||||:::      ::::::||||||   |||   |||||||||   |||   |||
Db   3437 GCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAA 3496
```

```
Qy      407 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 426
            |||      |||:::|||  |||||||||:::   |||   ||||||   |||   |||
Db     3497 GGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTAT 3556

Qy      427 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 446
            |||   |||||||||:::   :::|||:::|||||||||||||   :::::::::|||:::
Db     3557 CCCACCCGACCGGACATCCCAGTCCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAG 3616

Qy      447 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 466
            |||:::|||   |||||||||||||||||||||||||||||   ||||||||||||:::
Db     3617 ACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAG 3676

Qy      467 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 486
            |||   |||||||||:::   |||:::||||||   ||||||   :::   |||   |||
Db     3677 CGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTG 3736

Qy      487 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 506
            ||||||||||||||||||:::   :::   ||||||   |||||||||:::||||||:::
Db     3737 AATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGAC 3796

Qy      507 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 526
            ||||||||||||||||||||||||||||||   |||||||||||||||      :::|||
Db     3797 TGCAGCATTGCTGAGAACATTGCCTATGGACACAACAGCCGGGTGGTGTCACAGGAAGAG 3856

Qy      527 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 546
            |||         ||||||      ||||||||||||:::|||||||||   ||||||   |||
Db     3857 ATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAA 3916

Qy      547 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 566
            |||:::|||:::||||||   ||||||   |||||||||||||||||||||||||||:::
Db     3917 TATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATT 3976

Qy      567 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 586
            ||||||||||||||||||:::::::::|||   ||||||||||||||||||||||||||
Db     3977 GCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCA 4036

Qy      587 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 606
            |||||||||   :::||||||||||||||||||   ||||||||||||   |||   |||
Db     4037 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAACCCAGAGAAGGC 4096

Qy      607 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 626
            |||||||||:::|||:::   ||||||||||||   ||||||||||||||||||||||||
Db     4097 CGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTG 4156

Qy      627 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 646
            |||         ||||||::::::|||:::   ||||||||||||:::||||||      :::
Db     4157 GTGTTTCAGAATGGCAGAGTCAAGCAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAA 4216

Qy      647 AspIleTyrPheLysLeuValAsnAlaGln 656
               |||||||||   :::|||:::   |||
Db     4217 GGCATCTATTTTCAATGGTCAGTGTCCAG 4246
```

RESULT 13
US-08-996-545-1

```
; Sequence 1, Application US/08996545
; Patent No. 5928898
;  GENERAL INFORMATION:
;    APPLICANT:  Skatrud, Paul L.
;    APPLICANT:  de Waard, Maarten A.
;    APPLICANT:  Peery, Robert B.
;    APPLICANT:  Andrade, Alan C.
;    TITLE OF INVENTION:  Multiple Drug Resistance Gene atrD of
;    TITLE OF INVENTION:  Aspergillus Nidulans
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  U.S.
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/996,545
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Webster, Thomas D.
;      REGISTRATION NUMBER:  39,872
;      REFERENCE/DOCKET NUMBER:  X-11766
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-3334
;      TELEFAX:  317-276-2763
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4002 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..4002
US-08-996-545-1
```

```
Alignment Scores:
Pred. No.:               4.67e-143     Length:         4002
Score:                   1254.00       Matches:        274
Percent Similarity:      58.47%        Conservative:   123
Best Local Similarity:   40.35%        Mismatches:     260
Query Match:             37.74%        Indels:         22
DB:                      2             Gaps:           6
```

US-09-873-409-1 (1-659) x US-08-996-545-1 (1-4002)

```
Qy        2 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 21
            :::||||||:::|||   |||   ||||||:::   :::   |||   |||      |||
Db     1954 ATTGCTGAACAAGGAACTCACGATGAATTGGTTGACCGCGGAGGCGCTTATCGCAAACTT 2013

Qy       22 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 41
            |||   :::|||   |||   :::         |||
Db     2014 GTGGAGGCTCAACGTATCAATGAACAGAAGGAAGCTGACGCCTTGGAGGACGCCGACGCT 2073

Qy       42 GluArgLysThrAsnSer------LeuProLeuHisSerValLysSerIleLysSerAsp 59
            |||      ||||||:::         :::   :::      |||      ||||||
Db     2074 GAGGATCTCACGAATGCAGATATTGCCAAAATCAAAACTGCGTCAAGCGCATCATCCGAT 2133

Qy       60 Phe------------------IleAspLys------AlaGluGluSerThrGlnSerLys 71
                                 ||||||:::             :::|||      |||
Db     2134 CTCGACGGAAAACCCACAACCATTGACCGCACGGGCACCCACAAGTCTGTTTCCAGCGCG 2193

Qy       72 GluIleSer------------------LeuProGluValSerLeuLeuLysIleLeuLys 85
               :::|||                  |||:::   ||||||      :::||||||
Db     2194 ATTCTTTCTAAAAGACCCCCCGAAACAACTCCGAAATACTCATTATGGACGCTGCTCAAA 2253

Qy       86 ------------LeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 101
                        |||:::||||||   |||:::::::::::::|||   :::   |||
Db     2254 TTTGTTGCTTCCTTCAACCGCCCTGAAATCCCGTACATGCTCATCGGTCTTGTCTTCTCA 2313

Qy      102 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 121
            ||||||   |||         |||      ::::::::::::||||||   |||:::
Db     2314 GTGTTAGCTGGTGGTGGCCAACCCACGCAAGCAGTGCTATATGCTAAAGCCATCAGCACA 2373

Qy      122 PheGly------AsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 139
                        :::      :::   |||:::|||||||||:::   :::|||:::
Db     2374 CTCTCGCTCCCAGAATCACAATATAGCAAGCTTCGACATGATGCGGATTTCTGGTCATTG 2433

Qy      140 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 159
            :::|||   ::::::|||:::|||   |||::::::               |||      :::
```

```
Db   2434 ATGTTCTTCGTGGTTGGTATCATTCAGTTTATCACGCAGTCAACCAATGGTGCTGCATTT 2493

Qy    160 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 179
                       |||   |||        |||   |||      ||||||:::   :::|||
Db   2494 GCCGTATGCTCCGAGAGACTTATTCGTCGCGCGAGAAGCACTGCCTTTCGGACGATACTC 2553

Qy    180 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIle 199
             ||||||||||||:::||||||::::::||||||||||||||||   ||||||:::
Db   2554 CGTCAAGACATTGCTTTCTTTGACAAGGAAGAGAATAGCACCGGCGCTCTGACCTCTTTC 2613

Qy    200 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 219
          |||:::   :::         :::   |||   :::|||      :::|||   :::
Db   2614 CTGTCCACCGAGACGAAGCATCTCTCCGGTGTTAGCGGTGTGACTCTAGGCACGATCTTG 2673

Qy    220 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 239
                :::|||   :::|||   ::::::||||||:::         ||||||::::::
Db   2674 ATGACCTCCACGACCCTAGGAGCGGCTATCATTATTGCCCTGGCGATTGGGTGGAAATTG 2733

Qy    240 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 259
                :::   :::|||:::   |||||||||            |||
Db   2734 GCCTTAGTTTGTATCTCGGTTGTGCCGGTTCTCCTGGCATGCGGTTTCTACCGATTCTAT 2793

Qy    260 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 279
             :::      |||   ::::::   |||         :::   :::         |||
Db   2794 ATGCTAGCCCAGTTTCAATCACGCTCCAAGCTTGCTTATGAGGGATCTGCAAACTTTGCT 2853

Qy    280 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 299
             ||||||            :::|||||||||:::   ||||||||||||||||:::
Db   2854 TGCGAGGCTACATCGTCTATCCGCACAGTTGCGTCATTAACCCGGGAAAGGGATGTCTGG 2913

Qy    300 GlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 319
          ::::::|||            |||      |||   |||   :::
Db   2914 GAGATTTACCATGCCCAGCTTGACGCACAAGGCAGGACCAGTCTAATCTCTGTCTTGAGG 2973

Qy    320 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 339
                      ||||||   |||   |||   ::::::|||      |||   ||||||
Db   2974 TCATCCCTGTTATATGCGTCGTCGCAGGCACTTGTTTCTTCTGCGTTGCGCTCGGGTTT 3033

Qy    340 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 359
             :::|||      |||:::                                 |||:::
Db   3034 TGGTACGGAGGGACACTTCTTGGTCACCACGAGTATGACATTTTCCGCTTCTTTGTTTGT 3093

Qy    360 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 379
          |||:::   |||   :::||||||   :::   |||         ||||||:::
Db   3094 TTCTCCGAGATTCTCTTTGGTGCTCAATCCGCGGGCACCGTCTTTTCCTTTGCACCAGAC 3153

Qy    380 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 399
                |||||||||:::   ||||||            |||   ::::::||||||
Db   3154 ATGGGCAAGGCGAAGAATGCGGCCGCCGAATTCCGACGACTGTTCGACCGAAAGCCACAA 3213

Qy    400 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 419
          ||||||:::   |||:::||||||:::|||   :::|||   ||||||   :::||||||
Db   3214 ATTGATAACTGGTCTGAAGAGGGCGAGAAGCTCGAAACGGTGGAAGGTGAAATCGAATTT 3273

Qy    420 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 439
          |||   |||   |||   ||||||   ||||||:::      :::||||||||||||
Db   3274 AGGAACGTGCACTTCAGATACCCGACCCGCCCAGAACAGCCTGTCCTGCGCGGCTTGGAC 3333

Qy    440 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 459
          |||:::::::::   |||:::   ||||||      ||||||   ||||||||||||||||
Db   3334 CTGACCGTGAAGCCTGGACAATATGTTGCGCTTGTCGGACCCAGCGGTTGTGGCAAGAGT 3393

Qy    460 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 479
          |||::::::   |||||||:::|||   ||||||   :::   |||   :::|||   |||
Db   3394 ACCACCATTGCATTGCTTGAGCGCTTTTACGATGCGATTGCCGGGTCCATCCTTGTTGAT 3453

Qy    480 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 499
          |||   |||      :::||||||:::         ||||||   :::::::::|||
Db   3454 GGGAAGGACATAAGTAAACTAAATATCAACTCCTACCGCAGCTTTCTGTCACTGGTCAGC 3513

Qy    500 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 519
          |||||||||   |||:::      :::|||   |||||||||         |||
Db   3514 CAGGAGCCGACACTGTACCAGGGCACCATCAAGGAAAACATCTTACTTGGTATTGTCGAA 3573

Qy    520 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 539
                ||||||   :::   :::   :::|||         |||||||||:::   |||
Db   3574 GATGACGTACCGGAAGAATTCTTGATTAAGGCTTGCAAGGACGCTAATATCTACGACTTC 3633

Qy    540 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 559
          |||      |||||||||   :::||||||   ||||||   ||||||      ||||||
Db   3634 ATCATGTCGCTCCCGGAGGGCTTTAATACAGTTGTTGGCAGCAAGGGAGGCATGTTGTCT 3693

Qy    560 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 579
          ||||||||||||||||||:::|||||||||||||||||||||:::   ||||||||||||
Db   3694 GGCGGCCAAAAGCAACGTGTGGCCATTGCCCGAGCCCTTCTTCGGGATCCCAAAATCCTT 3753

Qy    580 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 599
          |||||||||||||||||||||||||||||||:::::::||||||||||||||||   |||
Db   3754 CTTCTCGATGAAGCGACGTCAGCCCTCGACTCCGAGTCAGAAAAGGTCGTCCAGGCGGCT 3813

Qy    600 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 619
          ||||||   |||      |||||||||   :::   |||   ||||||||||||   |||
Db   3814 TTGGATGCCGCTGCCCGAGGCCGAACCACAATCGCCGTTGCACACCGACTCAGCACGATT 3873

Qy    620 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 639
          |||   ||||||:::|||   |||         |||||||||   |||   |||||||||
Db   3874 CAAAAGGCGGACGTTATCTATGTTTTCGACCAAGGCAAGATCGTCGAAAGCGGAACGCAC 3933

Qy    640 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerVal 658
             ||||||::::::   :::      |||::::::|||||||||   ||||||:::
Db   3934 AGCGAACTGGTCCAGAAAAAGGGCCGGTACTACGAGCTGGTCAACTTGCAGAGCTTG 3990
```

RESULT 14
US-08-996-545-3
; Sequence 3, Application US/08996545
; Patent No. 5928898
; GENERAL INFORMATION:
;   APPLICANT: Skatrud, Paul L.
;   APPLICANT: de Waard, Maarten A.
;   APPLICANT: Peery, Robert B.
;   APPLICANT: Andrade, Alan C.
;   TITLE OF INVENTION: Multiple Drug Resistance Gene atrD of
;   TITLE OF INVENTION: Aspergillus Nidulans
;   NUMBER OF SEQUENCES: 3
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Eli Lilly and Company
;   STREET: Lilly Corporate Center
;   CITY: Indianapolis
;   STATE: Indiana
;   COUNTRY: U.S.
;   ZIP: 46285
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/996,545
;   FILING DATE:
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Webster, Thomas D.
;   REGISTRATION NUMBER: 39,872
;   REFERENCE/DOCKET NUMBER: X-11766
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 317-276-3334
;   TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4002 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: mRNA
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO

```
US-08-996-545-3

Alignment Scores:
Pred. No.:              4.67e-143     Length:         4002
Score:                  1254.00       Matches:        274
Percent Similarity:     58.47%        Conservative:   123
Best Local Similarity:  40.35%        Mismatches:     260
Query Match:            37.74%        Indels:         22
DB:                     2             Gaps:           6

US-09-873-409-1 (1-659) x US-08-996-545-3 (1-4002)

Qy      2 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 21
Db   1954 AUUGCUGAACAAGGAACUCACGAUGAAUUGGUUGACCGCGGAGGCGCUUAUCGCAAACUU 2013

Qy     22 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 41
Db   2014 GUGGAGGCUCAACGUAUCAAUGAACAGAAGGAAGCUGACGCCUUGGAGGACGCCGACGCU 2073

Qy     42 GluArgLysThrAsnSer------LeuProLeuHisSerValLysSerIleLysSerAsp 59
Db   2074 GAGGAUCUCACGAAUGCAGAUAUUGCCAAAAUCAAAACUGCGUCAAGCGCAUCAUCCGAU 2133

Qy     60 Phe------------------IleAspLys------AlaGluGluSerThrGlnSerLys 71
Db   2134 CUCGACGGAAAACCCACAACCAUUGACCGCACGGGCACCCACAAGUCUGUUUCCAGCGCG 2193

Qy     72 GluIleSer------------------LeuProGluValSerLeuLeuLysIleLeuLys 85
Db   2194 AUUCUUUCUAAAAGACCCCCCGAAACAACUCCGAAAUACUCAUUAUGGACGCUGCUCAAA 2253

Qy     86 ------------LeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 101
Db   2254 UUUGUUGCUUCCUUCAACCGCCCUGAAAUCCCGUACAUGCUCAUCGGUCUUGUCUUCUCA 2313

Qy    102 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 121
Db   2314 GUGUUAGCUGGUGGUGGCCAACCCACGCAAGCAGUGCUAUAUGCUAAAGCCAUCAGCACA 2373

Qy    122 PheGly------AsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 139
Db   2374 CUCUCGCUCCCAGAAUCACAAUAUAGCAAGCUUCGACAUGAUGCGGAUUUCUGGUCAUUG 2433

Qy    140 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 159
Db   2434 AUGUUCUUCGUGGUUGGUAUCAUUCAGUUUAUCACGCAGUCAACCAAUGGUGCUGCAUUU 2493

Qy    160 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 179
Db   2494 GCCGUAUGCUCCGAGAGACUUAUUCGUCGCGCGAGAAGCACUGCCUUUCGGACGAUACUC 2553

Qy    180 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIle 199
Db   2554 CGUCAAGACAUUGCUUUCUUUGACAAGGAAGAGAAUAGCACCGGCGCUCUGACCUCUUUC 2613

Qy    200 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 219
Db   2614 CUGUCCACCGAGACGAAGCAUCUCUCCGGUGUUAGCGGUGUGACUCUAGGCACGAUCUUG 2673

Qy    220 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 239
Db   2674 AUGACCUCCACGACCCUAGGAGCGGCUAUCAUUAUUGCCCUGGCGAUUGGGUGGAAAUUG 2733

Qy    240 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 259
Db   2734 GCCUUAGUUUGUAUCUCGGUUGUGCCGGUUCUCCUGGCAUGCGGUUUCUACCGAUUCUAU 2793

Qy    260 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 279
Db   2794 AUGCUAGCCCAGUUUCAAUCACGCUCCAAGCUUGCUUAUGAGGGAUCUGCAAACUUUGCU 2853

Qy    280 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 299
Db   2854 UGCGAGGCUACAUCGUCUAUCCGCACAGUUGCGUCAUUAACCCGGGAAAGGGAUGUCUGG 2913

Qy    300 GlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 319
Db   2914 GAGAUUUACCAUGCCCAGCUUGACGCACAAGGCAGGACCAGUCUAAUCUCUGUCUUGAGG 2973

Qy    320 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 339
Db   2974 UCAUCCCUGUUAUAUGCGUCGUCGCAGGCACUUGUUUUCUUCUGCGUUGCGCUCGGGUUU 3033

Qy    340 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 359
Db   3034 UGGUACGGAGGGACACUUCUUGGUCACCACGAGUAUGACAUUUUCCGCUUCUUUGUUUGU 3093

Qy    360 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 379
Db   3094 UUCUCCGAGAUUCUCUUUGGUGCUCAAUCCGCGGGCACCGUCUUUUCCUUUGCACCAGAC 3153

Qy    380 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 399
Db   3154 AUGGGCAAGGCGAAGAAUGCGGCCGCCGAAUUCCGACGACUGUUCGACCGAAAGCCACAA 3213

Qy    400 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 419
Db   3214 AUUGAUAACUGGUCUGAAGAGGGCGAGAAGCUCGAAACGGUGGAAGGUGAAAUCGAAUUU 3273

Qy    420 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 439
Db   3274 AGGAACGUGCACUUCAGAUACCCGACCCGCCCAGAACAGCCUGUCCUGCGCGGCUUGGAC 3333

Qy    440 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 459
Db   3334 CUGACCGUGAAGCCUGGACAAUAUGUUGCGCUUGUCGGACCCAGCGGUUGUGGCAAGAGU 3393

Qy    460 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 479
Db   3394 ACCACCAUUGCAUUGCUUGAGCGCUUUUACGAUGCGAUUGCCGGGUCCAUCCUUGUUGAU 3453

Qy    480 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 499
Db   3454 GGGAAGGACAUAAGUAAACUAAAUAUCAACUCCUACCGCAGCUUUCUGUCACUGGUCAGC 3513

Qy    500 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 519
Db   3514 CAGGAGCCGACACUGUACCAGGGCACCAUCAAGGAAAACAUCUUACUUGGUAUUGUCGAA 3573

Qy    520 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 539
Db   3574 GAUGACGUACCGGAAGAAUUCUUGAUUAAGGCUUGCAAGGACGCUAAUAUCUACGACUUC 3633

Qy    540 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 559
Db   3634 AUCAUGUCGCUCCCGGAGGGCUUUAAUACAGUUGUUGGCAGCAAGGGAGGCAUGUUGUCU 3693

Qy    560 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 579
Db   3694 GGCGGCCAAAAGCAACGUGUGGCCAUUGCCCGAGCCCUUCUUCGGGAUCCCAAAAUCCUU 3753

Qy    580 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 599
Db   3754 CUUCUCGAUGAAGCGACGUCAGCCCUCGACUCCGAGUCAGAAAAGGUCGUCCAGGCGGCU 3813

Qy    600 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 619
Db   3814 UUGGAUGCCGCUGCCCGAGGCCGAACCACAAUCGCCGUUGCACACCGACUCAGCACGAUU 3873

Qy    620 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 639
Db   3874 CAAAAGGCGGACGUUAUCUAUGUUUUCGACCAAGGCAAGAUCGUCGAAAGCGGAACGCAC 3933

Qy    640 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerVal 658
```

```
Db    3934 AGCGAACUGGUCCAGAAAAAGGGCCGGUACUACGAGCUGGUCAACUUGCAGAGCUUG 3990

RESULT 15
US-09-328-320-1
; Sequence 1, Application US/09328320
; Patent No. 6228615
;  GENERAL INFORMATION:
;    APPLICANT:  Skatrud, Paul L.
;    APPLICANT:  de Waard, Maarten A.
;    APPLICANT:  Peery, Robert B.
;    APPLICANT:  Andrade, Alan C.
;    TITLE OF INVENTION:  Multiple Drug Resistance Gene atrD of
;    TITLE OF INVENTION:  Aspergillus Nidulans
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  U.S.
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/328,320
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/996,545
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Webster, Thomas D.
;      REGISTRATION NUMBER:  39,872
;      REFERENCE/DOCKET NUMBER:  X-11766
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-3334
;      TELEFAX:  317-276-2763
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4002 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..4002
US-09-328-320-1

Alignment Scores:
Pred. No.:               4.67e-143      Length:        4002
Score:                   1254.00        Matches:       274
Percent Similarity:      58.47%         Conservative:  123
Best Local Similarity:   40.35%         Mismatches:    260
Query Match:             37.74%         Indels:        22
DB:                      4              Gaps:          6

US-09-873-409-1 (1-659) x US-09-328-320-1 (1-4002)

Qy       2 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 21
             :::||||||:::|||   |||   ||||||:::   :::   |||   |||     |||
Db    1954 ATTGCTGAACAAGGAACTCACGATGAATTGGTTGACCGCGGAGGCGCTTATCGCAAACTT 2013

Qy      22 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 41
           |||   :::|||   |||   :::         |||
Db    2014 GTGGAGGCTCAACGTATCAATGAACAGAAGGAAGCTGACGCCTTGGAGGACGCCGACGCT 2073

Qy      42 GluArgLysThrAsnSer------LeuProLeuHisSerValLysSerIleLysSerAsp 59
               |||      ||||||:::            :::   :::      |||      ||||||
Db    2074 GAGGATCTCACGAATGCAGATATTGCCAAAATCAAAACTGCGTCAAGCGCATCATCCGAT 2133

Qy      60 Phe------------------IleAspLys------AlaGluGluSerThrGlnSerLys 71
                                ||||||:::            :::|||      |||
Db    2134 CTCGACGGAAAACCCACAACCATTGACCGCACGGGCACCCACAAGTCTGTTTCCAGCGCG 2193

Qy      72 GluIleSer------------------LeuProGluValSerLeuLeuLysIleLeuLys 85
             :::|||                    |||:::   ||||||      :::||||||
Db    2194 ATTCTTTCTAAAAGACCCCCCGAAACAACTCCGAAATACTCATTATGGACGCTGCTCAAA 2253

Qy      86 ------------LeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 101
                       |||:::||||||   |||::::::::::::|||   :::   |||
Db    2254 TTTGTTGCTTCCTTCAACCGCCCTGAAATCCCGTACATGCTCATCGGTCTTGTCTTCTCA 2313

Qy     102 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 121
           ||||||   |||         |||      ::::::::::::||||||   |||:::
Db    2314 GTGTTAGCTGGTGGTGGCCAACCCACGCAAGCAGTGCTATATGCTAAAGCCATCAGCACA 2373

Qy     122 PheGly------AsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 139
                       :::      :::   |||:::|||||||||:::   :::|||:::
Db    2374 CTCTCGCTCCCAGAATCACAATATAGCAAGCTTCGACATGATGCGGATTTCTGGTCATTG 2433

Qy     140 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 159
           :::|||   ::::::|||:::|||   |||::::::            |||      :::
Db    2434 ATGTTCTTCGTGGTTGGTATCATTCAGTTTATCACGCAGTCAACCAATGGTGCTGCATTT 2493

Qy     160 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 179
                       |||   |||         |||   |||      ||||||:::   :::|||
Db    2494 GCCGTATGCTCCGAGAGACTTATTCGTCGCGCGAGAAGCACTGCCTTTCGGACGATACTC 2553

Qy     180 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIle 199
             ||||||||||||:::||||||::::::|||||||||||||||   ||||||:::
Db    2554 CGTCAAGACATTGCTTTCTTTGACAAGGAAGAGAATAGCACCGGCGCTCTGACCTCTTTC 2613

Qy     200 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 219
           |||:::   :::         :::   |||   :::|||      :::|||   :::
Db    2614 CTGTCCACCGAGACGAAGCATCTCTCCGGTGTTAGCGGTGTGACTCTAGGCACGATCTTG 2673

Qy     220 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 239
                 :::|||   :::|||   ::::::||||||            ||||||::::::
Db    2674 ATGACCTCCACGACCCTAGGAGCGGCTATCATTATTGCCCTGGCGATTGGGTGGAAATTG 2733

Qy     240 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 259
                 :::   :::|||:::   ||||||||         |||
Db    2734 GCCTTAGTTTGTATCTCGGTTGTGCCGGTTCTCCTGGCATGCGGTTTCTACCGATTCTAT 2793

Qy     260 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 279
             :::      |||   ::::::   |||            :::   :::         |||
Db    2794 ATGCTAGCCCAGTTTCAATCACGCTCCAAGCTTGCTTATGAGGGATCTGCAAACTTTGCT 2853

Qy     280 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 299
             ||||||      :::||||||||||:::   |||||||||||||||:::
Db    2854 TGCGAGGCTACATCGTCTATCCGCACAGTTGCGTCATTAACCCGGGAAAGGGATGTCTGG 2913

Qy     300 GlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 319
           ::::::|||         |||      |||   |||   :::
Db    2914 GAGATTTACCATGCCCAGCTTGACGCACAAGGCAGGACCAGTCTAATCTCTGTCTTGAGG 2973

Qy     320 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 339
                       ||||||   |||   |||   ::::::|||      |||   ||||||
Db    2974 TCATCCCTGTTATATGCGTCGTCGCAGGCACTTGTTTTCTTCTGCGTTGCGCTCGGGTTT 3033

Qy     340 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 359
             :::|||      |||:::                              |||:::
Db    3034 TGGTACGGAGGGACACTTCTTGGTCACCACGAGTATGACATTTTCCGCTTCTTTGTTTGT 3093

Qy     360 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 379
           |||:::   |||   :::||||||   :::   |||               ||||||:::
Db    3094 TTCTCCGAGATTCTCTTTGGTGCTCAATCCGCGGGCACCGTCTTTTCCTTTGCACCAGAC 3153

Qy     380 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 399
               |||||||||:::   ||||||               |||   ::::::||||||
```

```
Db   3154 ATGGGCAAGGCGAAGAATGCGGCCGCCGAATTCCGACGACTGTTCGACCGAAAGCCACAA 3213

Qy    400 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 419
          ||||||:::   |||:::||||||:::|||   :::|||   ||||||   :::||||||
Db   3214 ATTGATAACTGGTCTGAAGAGGGCGAGAAGCTCGAAACGGTGGAAGGTGAAATCGAATTT 3273

Qy    420 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 439
          |||   |||   |||   ||||||   ||||||:::         :::||||||||||||
Db   3274 AGGAACGTGCACTTCAGATACCCGACCCGCCCAGAACAGCCTGTCCTGCGCGGCTTGGAC 3333

Qy    440 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 459
          |||:::::::::   |||:::   ||||||   ||||||   ||||||||||||||||||
Db   3334 CTGACCGTGAAGCCTGGACAATATGTTGCGCTTGTCGGACCCAGCGGTTGTGGCAAGAGT 3393

Qy    460 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 479
          |||::::::   ||||||:::|||   ||||||   :::      |||   :::|||   |||
Db   3394 ACCACCATTGCATTGCTTGAGCGCTTTTACGATGCGATTGCCGGGTCCATCCTTGTTGAT 3453

Qy    480 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 499
          |||   |||         :::||||||:::               ||||||   :::::::::|||
Db   3454 GGGAAGGACATAAGTAAACTAAATATCAACTCCTACCGCAGCTTTCTGTCACTGGTCAGC 3513

Qy    500 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 519
          |||||||||   |||:::      :::|||   |||||||||         |||
Db   3514 CAGGAGCCGACACTGTACCAGGGCACCATCAAGGAAAACATCTTACTTGGTATTGTCGAA 3573

Qy    520 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 539
                ||||||   :::   :::   :::|||         |||||||||:::   |||
Db   3574 GATGACGTACCGGAAGAATTCTTGATTAAGGCTTGCAAGGACGCTAATATCTACGACTTC 3633

Qy    540 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 559
          |||      |||||||||   :::||||||   ||||||   ||||||      ||||||
Db   3634 ATCATGTCGCTCCCGGAGGGCTTTAATACAGTTGTTGGCAGCAAGGGAGGCATGTTGTCT 3693

Qy    560 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 579
          ||||||||||||||||||:::|||||||||||||||||||||:::   ||||||||||||
Db   3694 GGCGGCCAAAAGCAACGTGTGGCCATTGCCCGAGCCCTTCTTCGGGATCCCAAAATCCTT 3753

Qy    580 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 599
          ||||||||||||||||||||||||||||||::::::||||||||||||||||||   |||
Db   3754 CTTCTCGATGAAGCGACGTCAGCCCTCGACTCCGAGTCAGAAAAGGTCGTCCAGGCGGCT 3813

Qy    600 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 619
          ||||||   |||      |||||||||   :::   |||   ||||||||||||   |||
Db   3814 TTGGATGCCGCTGCCCGAGGCCGAACCACAATCGCCGTTGCACACCGACTCAGCACGATT 3873

Qy    620 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 639
          |||   ||||||:::|||   |||      |||||||||   |||   |||||||||
Db   3874 CAAAAGGCGGACGTTATCTATGTTTTCGACCAAGGCAAGATCGTCGAAAGCGGAACGCAC 3933

Qy    640 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerVal 658
             ||||||::::::   :::      |||::::::|||||||||   ||||||:::
Db   3934 AGCGAACTGGTCCAGAAAAAGGGCCGGTACTACGAGCTGGTCAACTTGCAGAGCTTG 3990
```

Search completed: March 31, 2003, 15:20:38
Job time : 192.507 secs

```
                GenCore version 5.1.4_p5_4578
           Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 31, 2003, 13:52:37 ; Search time 47.683 Seconds
                                           (without alignments)
                                           2337.888 Million cell updates/sec

Title:          US-09-873-409-3
Perfect score:  659
Sequence:       1 MVDENDIRALNVRHYRDHIG..........SESKSAVQAALEKDTPRYSF 131

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters:      1148742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115424_19290/app_query.fasta_1.7544
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09873409_@CGN_1_1_960_@runat_27032003_115424_19290
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Published_Applications_NA:*
                1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
                2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
                3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
                4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
                5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
                6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
                7:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
                8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
                9:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
                10:  /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
                11:  /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
                12:  /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
                13:  /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
                14:  /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	659	100.0	1175	10	US-09-873-409-11	Sequence 11, Appl
	2	659	100.0	1940	10	US-09-873-409-16	Sequence 16, Appl
	3	659	100.0	2021	10	US-09-873-409-15	Sequence 15, Appl
	4	619	93.9	2856	10	US-09-873-409-10	Sequence 10, Appl
	5	619	93.9	3177	10	US-09-873-409-12	Sequence 12, Appl
	6	619	93.9	3621	10	US-09-873-409-14	Sequence 14, Appl
	7	619	93.9	3702	10	US-09-873-409-13	Sequence 13, Appl
	~~8~~	~~468~~	~~71.0~~	~~4189~~	~~10~~	~~US-09-866-866A-5~~	~~Sequence 5, Appli~~
	9	467	70.9	4254	10	US-09-917-800A-1424	Sequence 1424, Ap
	10	467	70.9	4369	10	US-09-769-097-1	Sequence 1, Appli
	11	467	70.9	4425	10	US-09-769-097-3	Sequence 3, Appli
	12	465	70.6	4317	9	US-10-044-671-1	Sequence 1, Appli
	13	462	70.1	4788	10	US-09-866-866A-7	Sequence 7, Appli
	14	461	70.0	3860	10	US-09-866-866A-1	Sequence 1, Appli
	15	461	70.0	3860	10	US-09-866-866A-3	Sequence 3, Appli
	16	461	70.0	4643	9	US-10-072-621-2	Sequence 2, Appli
	17	461	70.0	8630	10	US-09-306-417-1	Sequence 1, Appli
	18	461	70.0	8630	10	US-09-306-417-2	Sequence 2, Appli
	19	448	68.0	3912	10	US-09-917-800A-1560	Sequence 1560, Ap
	20	440	66.8	3924	10	US-09-880-107-2299	Sequence 2299, Ap
	21	422	64.0	5010	10	US-09-917-800A-483	Sequence 483, App
	22	369	56.0	3512	10	US-09-749-340-7	Sequence 7, Appli
	23	368.5	55.9	1076	9	US-09-764-884-13	Sequence 13, Appl
	24	368.5	55.9	1076	9	US-10-092-256-13	Sequence 13, Appl
	25	368.5	55.9	1367	9	US-09-764-884-21	Sequence 21, Appl
	26	368.5	55.9	1367	9	US-10-092-256-21	Sequence 21, Appl
	27	366	55.5	4002	9	US-09-758-828-1	Sequence 1, Appli
	28	366	55.5	4002	9	US-09-758-828-3	Sequence 3, Appli
	29	364	55.2	1810	10	US-09-749-340-4	Sequence 4, Appli
	30	364	55.2	4175	10	US-09-749-340-3	Sequence 3, Appli
	31	363	55.1	3861	9	US-09-938-842A-263	Sequence 263, App
	32	357	54.2	4653	9	US-10-101-388-2	Sequence 2, Appli
c	33	356	54.0	981	10	US-09-770-445-272	Sequence 272, App
	34	350	53.1	216	10	US-09-864-761-27861	Sequence 27861, A
	35	350	53.1	465	10	US-09-864-761-11237	Sequence 11237, A
	36	347	52.7	939	10	US-09-974-300-6511	Sequence 6511, Ap
	37	339	51.4	2681	10	US-09-749-340-8	Sequence 8, Appli
	38	335	50.8	3861	9	US-09-938-842A-2237	Sequence 2237, Ap
	39	333	50.5	2698	10	US-09-749-340-5	Sequence 5, Appli
c	40	327	49.6	536165	9	US-09-939-964-1	Sequence 1, Appli
	41	326	49.5	8036	9	US-10-101-388-1	Sequence 1, Appli
	42	325	49.3	2066	10	US-09-873-409-9	Sequence 9, Appli
	43	319	48.4	2298	9	US-10-156-239-18	Sequence 18, Appl
	44	319	48.4	2298	10	US-09-795-693-18	Sequence 18, Appl
	45	319	48.4	3408	9	US-10-156-239-16	Sequence 16, Appl

ALIGNMENTS

```
RESULT 1
US-09-873-409-11
; Sequence 11, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 11
;   LENGTH: 1175
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-409-11

Alignment Scores:
Pred. No.:               1.58e-81      Length:        1175
Score:                   659.00        Matches:       131
Percent Similarity:      100.00%       Conservative:  0
Best Local Similarity:   100.00%       Mismatches:    0
Query Match:             100.00%       Indels:        0
DB:                      10            Gaps:          0
```

```
US-09-873-409-3 (1-131) x US-09-873-409-11 (1-1175)

Qy        1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      418 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 477

Qy       21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      478 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 537

Qy       41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      538 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 597

Qy       61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      598 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 657

Qy       81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      658 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 717

Qy      101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      718 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 777

Qy      121 AlaLeuGluLysAspThrProArgTyrSerPhe 131
              |||||||||||||||||||||||||||||||||
Db      778 GCACTGGAGAAGGATACCCCCAGGTATTCATTT 810

RESULT 2
US-09-873-409-16
; Sequence 16, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-409-16

Alignment Scores:
Pred. No.:              3.21e-81      Length:         1940
Score:                  659.00        Matches:        131
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            100.00%       Indels:         0
DB:                     10            Gaps:           0

US-09-873-409-3 (1-131) x US-09-873-409-16 (1-1940)

Qy        1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1183 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 1242

Qy       21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1243 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 1302

Qy       41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1303 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 1362

Qy       61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1363 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 1422

Qy       81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1423 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 1482

Qy      101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1483 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 1542

Qy      121 AlaLeuGluLysAspThrProArgTyrSerPhe 131
              |||||||||||||||||||||||||||||||||
Db     1543 GCACTGGAGAAGGATACCCCCAGGTATTCATTT 1575

RESULT 3
US-09-873-409-15
; Sequence 15, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
;   LENGTH: 2021
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: No. US20020037522A1e
;   LOCATION: (723)..(723)
;   OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-15

Alignment Scores:
Pred. No.:              3.4e-81       Length:         2021
Score:                  659.00        Matches:        131
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            100.00%       Indels:         0
DB:                     10            Gaps:           0

US-09-873-409-3 (1-131) x US-09-873-409-15 (1-2021)

Qy        1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1264 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 1323

Qy       21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1324 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 1383

Qy       41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1384 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 1443

Qy       61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1444 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 1503

Qy       81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1504 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 1563

Qy      101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1564 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 1623
```

```
Qy    121 AlaLeuGluLysAspThrProArgTyrSerPhe 131
          |||||||||||||||||||||||||||||||||
Db   1624 GCACTGGAGAAGGATACCCCCAGGTATTCATTT 1656

RESULT 4
US-09-873-409-10
; Sequence 10, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
;   LENGTH: 2856
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-409-10

Alignment Scores:
Pred. No.:              1.84e-75      Length:        2856
Score:                  619.00        Matches:       124
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            93.93%        Indels:        0
DB:                     10            Gaps:          0

US-09-873-409-3 (1-131) x US-09-873-409-10 (1-2856)

Qy      1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    418 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 477

Qy     21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    478 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 537

Qy     41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    538 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 597

Qy     61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    598 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 657

Qy     81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    658 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 717

Qy    101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    718 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 777

Qy    121 AlaLeuGluLys 124
          ||||||||||||
Db    778 GCACTGGAGAAG 789

RESULT 5
US-09-873-409-12
; Sequence 12, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
;   LENGTH: 3177
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: No. US20020037522A1e
;   LOCATION: (198)..(198)
;   OTHER INFORMATION: n at position 198 represents any nucleotide (A, T, C or G)
US-09-873-409-12

Alignment Scores:
Pred. No.:              2.14e-75      Length:        3177
Score:                  619.00        Matches:       124
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            93.93%        Indels:        0
DB:                     10            Gaps:          0

US-09-873-409-3 (1-131) x US-09-873-409-12 (1-3177)

Qy      1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    739 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 798

Qy     21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    799 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 858

Qy     41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    859 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 918

Qy     61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    919 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 978

Qy     81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    979 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 1038

Qy    101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1039 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 1098

Qy    121 AlaLeuGluLys 124
          ||||||||||||
Db   1099 GCACTGGAGAAG 1110

RESULT 6
US-09-873-409-14
; Sequence 14, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
;   LENGTH: 3621
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-409-14

Alignment Scores:
Pred. No.:              2.57e-75      Length:        3621
```

```
Score:                     619.00         Matches:        124
Percent Similarity:        100.00%        Conservative:   0
Best Local Similarity:     100.00%        Mismatches:     0
Query Match:               93.93%         Indels:         0
DB:                        10             Gaps:           0

US-09-873-409-3 (1-131) x US-09-873-409-14 (1-3621)

Qy      1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1183 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 1242

Qy     21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1243 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 1302

Qy     41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1303 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 1362

Qy     61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1363 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 1422

Qy     81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1423 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 1482

Qy    101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1483 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 1542

Qy    121 AlaLeuGluLys 124
          ||||||||||||
Db   1543 GCACTGGAGAAG 1554

RESULT 7
US-09-873-409-13
; Sequence 13, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
;   LENGTH: 3702
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: No. US20020037522A1e
;   LOCATION: (723)..(723)
;   OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-13

Alignment Scores:
Pred. No.:                 2.65e-75       Length:         3702
Score:                     619.00         Matches:        124
Percent Similarity:        100.00%        Conservative:   0
Best Local Similarity:     100.00%        Mismatches:     0
Query Match:               93.93%         Indels:         0
DB:                        10             Gaps:           0

US-09-873-409-3 (1-131) x US-09-873-409-13 (1-3702)

Qy      1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1264 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 1323

Qy     21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1324 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 1383

Qy     41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1384 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 1443

Qy     61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1444 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 1503

Qy     81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1504 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 1563

Qy    101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1564 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 1623

Qy    121 AlaLeuGluLys 124
          ||||||||||||
Db   1624 GCACTGGAGAAG 1635

RESULT 8
US-09-866-866A-5
; Sequence 5, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
;  APPLICANT: Sorrentino, Brian
;  APPLICANT:  Schuetz, John
;  TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
;  FILE REFERENCE: 1340-1-021CIP2
;  CURRENT APPLICATION NUMBER: US/09/866,866A
;  CURRENT FILING DATE:  2001-08-30
;  PRIOR APPLICATION NUMBER: 09/584,586
;  PRIOR FILING DATE: 2000-05-31
;  PRIOR APPLICATION NUMBER: PCT/US99/11825
;  PRIOR FILING DATE: 1999-05-27
;  PRIOR APPLICATION NUMBER: 60/086,988
;  PRIOR FILING DATE: 1998-05-28
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 4189
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-866-866A-5

Alignment Scores:
Pred. No.:                 2.2e-54        Length:         4189
Score:                     468.00         Matches:        90
Percent Similarity:        91.06%         Conservative:   22
Best Local Similarity:     73.17%         Mismatches:     11
Query Match:               71.02%         Indels:         0
DB:                        10             Gaps:           0

US-09-873-409-3 (1-131) x US-09-866-866A-5 (1-4189)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
          :::|||      ||||||||   :::|||||||||:::   |||:::   |||||||||
Db   1354 ATCGACGGACAAGACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCATTGGTGTG 1413

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
          ||||||||||||||||||||||||   |||||||||:::   ||||||:::|||||||||
Db   1414 GTGAGTCAGGAACCTGTGCTGTTTGCCACCACGATCGCCGAGAACATTCGCTATGGCCGA 1473

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
          :::|||||||||   :::|||:::|||:::|||   :::||||||||||||||||||||||
Db   1474 GAAGATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTC 1533

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
```

```
               ||||||:::   |||::::::|||:::|||||||||||||:::|||||||||:::|||
Db      1534 ATCATGAAACTGCCCCACCAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGT 1593

Qy        82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1594 GGGGACAGAAACAGAGAATCGCCATTGCCCGGGCCCTGGTCCGCAATCCCAAGATCCTT 1653

Qy       102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
             :::|||||||||||||||||||||||||||:::||||||::::::   ||||||||||||
Db      1654 TTGTTGGACGAGGCCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTGGTGCAGGCCGCA 1713

Qy       122 LeuGluLys 124
             |||:::|||
Db      1714 CTGGATAAG 1722

RESULT 9
US-09-917-800A-1424
; Sequence 1424, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
;  APPLICANT: Mendrick, Donna
;  APPLICANT:  Porter, Mark
;  APPLICANT:  Johnson, Kory
;  APPLICANT:  Castle, Arthur
;  APPLICANT:  Elashoff, Michael
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Molecular Toxicology Modeling
;  FILE REFERENCE: 44921-5038-US
;  CURRENT APPLICATION NUMBER: US/09/917,800A
;  CURRENT FILING DATE:  2001-07-31
;  PRIOR APPLICATION NUMBER: US 60/222,040
;  PRIOR FILING DATE: 2000-07-31
;  PRIOR APPLICATION NUMBER: US 60/222,880
;  PRIOR FILING DATE: 2000-11-02
;  PRIOR APPLICATION NUMBER: US 60/290,029
;  PRIOR FILING DATE: 2001-05-11
;  PRIOR APPLICATION NUMBER: US 60/290,645
;  PRIOR FILING DATE: 2001-05-15
;  PRIOR APPLICATION NUMBER: US 60/292,336
;  PRIOR FILING DATE: 2001-05-22
;  PRIOR APPLICATION NUMBER: US 60/295,798
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: US 60/297,457
;  PRIOR FILING DATE: 2001-06-13
;  PRIOR APPLICATION NUMBER: US 60/298,884
;  PRIOR FILING DATE: 2001-06-19
;  PRIOR APPLICATION NUMBER: US 60/303,459
;  PRIOR FILING DATE: 2001-07-09
;  NUMBER OF SEQ ID NOS: 1740
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1424
;   LENGTH: 4254
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020119462A1 M81855
US-09-917-800A-1424

Alignment Scores:
Pred. No.:               3.09e-54      Length:         4254
Score:                   467.00        Matches:        90
Percent Similarity:      91.06%        Conservative:   22
Best Local Similarity:   73.17%        Mismatches:     11
Query Match:             70.86%        Indels:         0
DB:                      10            Gaps:           0

US-09-873-409-3 (1-131) x US-09-917-800A-1424 (1-4254)

Qy         2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
             :::|||      |||||||||     :::|||||||||:::    |||:::    |||||||||
Db      1456 ATTGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTG 1515

Qy        22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
             ||||||||||||||||||||||   |||||||||:::    ||||||:::|||||||||
Db      1516 GTGAGTCAGGAACCCGTGCTGTTTGCCACCACGATTGCCGAAAACATTCGCTATGGCCGA 1575

Qy        42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
             ::::::||||||    :::|||:::|||:::|||    :::||||||||||||||||||||
Db      1576 GAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTC 1635

Qy        62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
             ||||||:::   |||:::|||:::||||||||||||||||:::|||||||||:::|||
Db      1636 ATCATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGT 1695

Qy        82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1696 GGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAACCCCAAGATCCTT 1755

Qy       102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
             :::|||||||||||||||||||||||||||:::||||||::::::   ||||||||||||
Db      1756 TTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTGGTTCAGGCCGCT 1815

Qy       122 LeuGluLys 124
             |||:::|||
Db      1816 CTGGATAAG 1824

RESULT 10
US-09-769-097-1
; Sequence 1, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
;  APPLICANT: Kimberly Anne Brun
;  APPLICANT:  Richard James Chenery
;  APPLICANT:  Harma Ellens
;  APPLICANT:  John Anthony Feild
;  APPLICANT:  Lin Yue
;  TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
;  TITLE OF INVENTION:  ENCODING RAT MDR1A AND SCREENING METHODS THEREOF
;  FILE REFERENCE: GP-50009-C2
;  CURRENT APPLICATION NUMBER: US/09/769,097
;  CURRENT FILING DATE:  2001-01-25
;  PRIOR APPLICATION NUMBER: 09/208,809
;  PRIOR FILING DATE: 1998-12-09
;  PRIOR APPLICATION NUMBER: 09/156,800
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: US99/20770
;  PRIOR FILING DATE: 1999-09-10
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 4369
;   TYPE: DNA
;   ORGANISM: RATTUS RATTUS
US-09-769-097-1

Alignment Scores:
Pred. No.:               3.21e-54      Length:         4369
Score:                   467.00        Matches:        90
Percent Similarity:      91.06%        Conservative:   22
Best Local Similarity:   73.17%        Mismatches:     11
Query Match:             70.86%        Indels:         0
DB:                      10            Gaps:           0

US-09-873-409-3 (1-131) x US-09-769-097-1 (1-4369)

Qy         2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
             :::|||      |||||||||     :::|||||||||:::    |||:::    |||||||||
Db      1684 ATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTG 1743

Qy        22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
             ||||||||||||||||||||||   |||||||||:::    ||||||:::|||||||||
Db      1744 GTGAGTCAGGAACCCGTGCTGTTTGCCACCACAATTGCCGAAAACATTCGCTATGGCCGA 1803

Qy        42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
             ::::::||||||    :::|||:::|||:::|||    :::||||||||||||||||||||
Db      1804 GAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAATGCCTATGATTTC 1863
```

```
Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
          ||||||:::   |||:::||||||:::||||||||||||||:::|||||||||:::|||
Db   1864 ATCATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGT 1923

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1924 GGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAACCCCAAGATCCTT 1983

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
          :::|||||||||||||||||||||||||:::||||||::::::   |||||||||||
Db   1984 TTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTGGTTCAGGCCGCT 2043

Qy    122 LeuGluLys 124
          |||:::|||
Db   2044 CTGGATAAG 2052
```

RESULT 11
US-09-769-097-3
; Sequence 3, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: ENCODING RAT MDR1A AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP-50009-C2
; CURRENT APPLICATION NUMBER: US/09/769,097
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 4425
; TYPE: DNA
; ORGANISM: RATTUS RATTUS
US-09-769-097-3

```
Alignment Scores:
Pred. No.:                3.27e-54      Length:         4425
Score:                    467.00        Matches:        90
Percent Similarity:       91.06%        Conservative:   22
Best Local Similarity:    73.17%        Mismatches:     11
Query Match:              70.86%        Indels:         0
DB:                       10            Gaps:           0

US-09-873-409-3 (1-131) x US-09-769-097-3 (1-4425)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
          :::|||         ||||||||   :::|||||||||:::    |||:::    ||||||||
Db   1684 ATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTG 1743

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
          ||||||||||||||||||||||||   |||||||||:::   ||||||:::|||||||||
Db   1744 GTGAGTCAGGAACCCGTGCTGTTTGCCACCACAATTGCCGAAAACATTCGCTATGGCCGA 1803

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
          ::::::||||||   :::|||:::|||:::|||   :::||||||||||||||||||||
Db   1804 GAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAATGCCTATGATTTC 1863

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
          ||||||:::   |||:::||||||:::||||||||||||||:::|||||||||:::|||
Db   1864 ATCATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGT 1923

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1924 GGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAACCCCAAGATCCTT 1983

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
          :::|||||||||||||||||||||||||:::||||||::::::   |||||||||||
Db   1984 TTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTGGTTCAGGCCGCT 2043

Qy    122 LeuGluLys 124
          |||:::|||
Db   2044 CTGGATAAG 2052
```

RESULT 12
US-10-044-671-1
; Sequence 1, Application US/10044671
; Patent No. US20020177147A1
; GENERAL INFORMATION:
; APPLICANT: Washington State University Research Foundation
; APPLICANT: Mealey, Katrina
; APPLICANT: Bentjen, Steven
; TITLE OF INVENTION: MDR1 VARIANTS AND METHODS FOR THEIR USE
; FILE REFERENCE: 4630-61733
; CURRENT APPLICATION NUMBER: US/10/044,671
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/261,578
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/314,829
; PRIOR FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4317
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (70)..(3912)
; OTHER INFORMATION:
US-10-044-671-1

```
Alignment Scores:
Pred. No.:                5.96e-54      Length:         4317
Score:                    465.00        Matches:        89
Percent Similarity:       89.43%        Conservative:   21
Best Local Similarity:    72.36%        Mismatches:     13
Query Match:              70.56%        Indels:         0
DB:                       9             Gaps:           0

US-09-873-409-3 (1-131) x US-10-044-671-1 (1-4317)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
          :::|||         ||||||||   :::|||||||||||    |||:::       ||||||
Db   1429 ATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTG 1488

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
          ||||||||||||||||||||||||   |||||||||:::   ||||||:::|||||||||
Db   1489 GTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGC 1548

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
          ::::::||||||   :::|||:::|||:::|||   :::||||||||||||||||||||
Db   1549 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTT 1608

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
          ||||||:::   ||||||||||||:::||||||||||||||:::|||||||::::::|||
Db   1609 ATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAGAGAGGGGCCCGGCTGAGT 1668

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1669 GGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTT 1728

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
          :::|||||||||||||||||||||||||:::||||||::::::   ||||||   |||
Db   1729 CTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCC 1788
```

```
Qy    122 LeuGluLys 124
          |||:::|||
Db   1789 CTGGATAAG 1797

RESULT 13
US-09-866-866A-7
; Sequence 7, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
;  APPLICANT: Sorrentino, Brian
;  APPLICANT:  Schuetz, John
;  TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
;  FILE REFERENCE: 1340-1-021CIP2
;  CURRENT APPLICATION NUMBER: US/09/866,866A
;  CURRENT FILING DATE:  2001-08-30
;  PRIOR APPLICATION NUMBER: 09/584,586
;  PRIOR FILING DATE: 2000-05-31
;  PRIOR APPLICATION NUMBER: PCT/US99/11825
;  PRIOR FILING DATE: 1999-05-27
;  PRIOR APPLICATION NUMBER: 60/086,988
;  PRIOR FILING DATE: 1998-05-28
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
;   LENGTH: 4788
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-866-866A-7

Alignment Scores:
Pred. No.:              1.79e-53       Length:        4788
Score:                  462.00         Matches:       89
Percent Similarity:     90.24%         Conservative:  22
Best Local Similarity:  72.36%         Mismatches:    12
Query Match:            70.11%         Indels:        0
DB:                     10             Gaps:          0

US-09-873-409-3 (1-131) x US-09-866-866A-7 (1-4788)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
          :::|||       ||||||||    :::|||||||||:::    |||:::    |||||||||
Db   1345 ATCGACGGACAGGACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCATTGGTGTG 1404

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
          ||||||||||||||||||||||||     |||||||||:::    ||||||:::||||||||
Db   1405 GTGAGTCAGGAACCTGTGCTGTTTGCCACCACGATCGCCGAGAACATTCGCTATGGCCGA 1464

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
          :::|||||||||    :::|||:::|||:::|||    :::||||||||||||||||||||
Db   1465 GAAGATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTC 1524

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
          |||||||:::    |||::::::|||:::|||||||||||||||:::||||||    :::|||
Db   1525 ATCATGAAACTGCCCCACCAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCACGTGAGT 1584

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1585 GGGGACAGAAACAGAGAATCGCCATTGCCCGGGCCCTGGTCCGCAATCCCAAGATCCTT 1644

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
          :::||||||||||||||||||||||||:::||||||::::::      |||||||||||
Db   1645 TTGTTGGACGAGGCCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTGGTTCAGGCCGCA 1704

Qy    122 LeuGluLys 124
          |||:::|||
Db   1705 CTGGATAAG 1713

RESULT 14
US-09-866-866A-1
; Sequence 1, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
;  APPLICANT: Sorrentino, Brian
;  APPLICANT:  Schuetz, John
;  TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
;  FILE REFERENCE: 1340-1-021CIP2
;  CURRENT APPLICATION NUMBER: US/09/866,866A
;  CURRENT FILING DATE:  2001-08-30
;  PRIOR APPLICATION NUMBER: 09/584,586
;  PRIOR FILING DATE: 2000-05-31
;  PRIOR APPLICATION NUMBER: PCT/US99/11825
;  PRIOR FILING DATE: 1999-05-27
;  PRIOR APPLICATION NUMBER: 60/086,988
;  PRIOR FILING DATE: 1998-05-28
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 3860
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-866A-1

Alignment Scores:
Pred. No.:              1.82e-53       Length:        3860
Score:                  461.00         Matches:       90
Percent Similarity:     89.43%         Conservative:  20
Best Local Similarity:  73.17%         Mismatches:    13
Query Match:            69.95%         Indels:        0
DB:                     10             Gaps:          0

US-09-873-409-3 (1-131) x US-09-866-866A-1 (1-3860)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
          ||||||       |||||||||    :::|||||||||    |||:::    |||||||||
Db   1357 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 1416

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
          ||||||||||||||||||||||||     |||||||||:::    ||||||:::||||||||
Db   1417 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 1476

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
          ::::::||||||    :::|||:::|||:::|||    :::||||||||||||||||||||
Db   1477 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1536

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
          |||||||:::    |||:::||||||:::|||||||||||||||:::||||||||:::|||
Db   1537 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 1596

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1597 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 1656

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
          :::||||||||||||||||||||||||:::||||||::::::      ||||||    |||
Db   1657 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 1716

Qy    122 LeuGluLys 124
          |||:::|||
Db   1717 CTGGATAAG 1725

RESULT 15
US-09-866-866A-3
; Sequence 3, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
;  APPLICANT: Sorrentino, Brian
;  APPLICANT:  Schuetz, John
;  TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
;  FILE REFERENCE: 1340-1-021CIP2
;  CURRENT APPLICATION NUMBER: US/09/866,866A
;  CURRENT FILING DATE:  2001-08-30
;  PRIOR APPLICATION NUMBER: 09/584,586
;  PRIOR FILING DATE: 2000-05-31
;  PRIOR APPLICATION NUMBER: PCT/US99/11825
;  PRIOR FILING DATE: 1999-05-27
;  PRIOR APPLICATION NUMBER: 60/086,988
```

```
;  PRIOR FILING DATE: 1998-05-28
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
;   LENGTH: 3860
;   TYPE: DNA
;   ORGANISM: homo sapiens
US-09-866-866A-3

Alignment Scores:
Pred. No.:              1.82e-53      Length:         3860
Score:                  461.00        Matches:        90
Percent Similarity:     89.43%        Conservative:   20
Best Local Similarity:  73.17%        Mismatches:     13
Query Match:            69.95%        Indels:         0
DB:                     10            Gaps:           0

US-09-873-409-3 (1-131) x US-09-866-866A-3 (1-3860)

Qy        2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
            ||||||      |||||||||   :::|||||||||      |||:::   |||||||||
Db     1357 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 1416

Qy       22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
            ||||||||||||||||||||||||   |||||||||:::   ||||||:::|||||||||
Db     1417 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 1476

Qy       42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
            ::::::||||||   :::|||:::|||:::|||   :::|||||||||||||||||||||
Db     1477 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1536

Qy       62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
            ||||||:::   |||:::||||||:::|||||||||||||||:::|||||||||:::|||
Db     1537 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 1596

Qy       82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1597 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 1656

Qy      102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
            :::|||||||||||||||||||||||||||:::||||||::::::   ||||||   |||
Db     1657 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 1716

Qy      122 LeuGluLys 124
            |||:::|||
Db     1717 CTGGATAAG 1725

Search completed: March 31, 2003, 16:00:29
Job time : 58.683 secs
```

```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 31, 2003, 02:24:20 ; Search time 19.1843 Seconds
                                           (without alignments)
                                           2094.145 Million cell updates/sec

Title:          US-09-873-409-3
Perfect score:  659
Sequence:       1 MVDENDIRALNVRHYRDHIG..........SESKSAVQAALEKDTPRYSF 131

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters:      882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115422_19259/app_query.fasta_1.7544
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_578_@runat_27032003_115422_19259 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Issued_Patents_NA:*
                1:  /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
                2:  /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
                3:  /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
                4:  /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
                5:  /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
                6:  /cgn2_6/ptodata/2/ina/backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	472	71.6	4233	3	US-09-120-513-1	Sequence 1, Appli
	2	472	71.6	4233	4	US-09-450-105-1	Sequence 1, Appli
	3	461	70.0	4264	2	US-08-784-649A-1	Sequence 1, Appli
	4	461	70.0	4264	2	US-08-784-649A-5	Sequence 5, Appli
	5	461	70.0	4646	1	US-08-181-471-2	Sequence 2, Appli
	6	461	70.0	4669	2	US-08-752-447-1	Sequence 1, Appli
	7	461	70.0	4669	4	US-09-316-167-1	Sequence 1, Appli
	8	461	70.0	4669	6	5206352-3	Patent No. 5206352
	9	461	70.0	6505	2	US-08-793-610-5	Sequence 5, Appli
	10	461	70.0	9318	2	US-08-793-610-6	Sequence 6, Appli
	11	449	68.1	4669	2	US-08-583-276-18	Sequence 18, Appl
	12	372	56.4	2726	1	US-08-461-823-1	Sequence 1, Appli
	13	366	55.5	4002	2	US-08-996-545-1	Sequence 1, Appli
	14	366	55.5	4002	2	US-08-996-545-3	Sequence 3, Appli
	15	366	55.5	4002	4	US-09-328-320-1	Sequence 1, Appli
	16	366	55.5	4002	4	US-09-328-320-3	Sequence 3, Appli
	17	360	54.6	3924	1	US-08-395-246C-1	Sequence 1, Appli
	18	360	54.6	4047	2	US-08-612-734B-1	Sequence 1, Appli
	19	355	53.9	4224	1	US-08-612-521-1	Sequence 1, Appli
	20	346.5	52.6	4800	2	US-08-612-734B-3	Sequence 3, Appli
	21	325.5	49.4	2376	1	US-08-394-880B-1	Sequence 1, Appli
	22	319	48.4	1749	4	US-09-134-001C-1893	Sequence 1893, Ap
	23	309	46.9	3924	2	US-08-996-644-3	Sequence 3, Appli
	24	309	46.9	3924	3	US-09-352-552-3	Sequence 3, Appli
	25	309	46.9	3927	2	US-08-996-644-1	Sequence 1, Appli
	26	309	46.9	3927	3	US-09-352-552-1	Sequence 1, Appli
	27	300.5	45.6	6143	1	US-08-612-521-3	Sequence 3, Appli
	28	276	41.9	2061	4	US-09-061-764A-17	Sequence 17, Appl
	29	275	41.7	6492	4	US-08-961-527-188	Sequence 188, App
c	30	272	41.3	3576	4	US-08-976-259-79	Sequence 79, Appl
	31	271	41.1	1743	4	US-09-134-001C-2774	Sequence 2774, Ap
	32	270	41.0	775	4	US-09-221-017B-860	Sequence 860, App
	33	270	41.0	1959	4	US-09-061-764A-4	Sequence 4, Appli
	34	268	40.7	7721	3	US-08-772-270A-14	Sequence 14, Appl
	35	267.5	40.6	3792	4	US-09-351-224E-10	Sequence 10, Appl
	36	266	40.4	5120	3	US-08-772-270A-6	Sequence 6, Appli
	37	266	40.4	8370	2	US-08-488-706-1	Sequence 1, Appli
	38	265	40.2	7760	4	US-08-961-527-63	Sequence 63, Appl
c	39	265	40.2	28804	2	US-08-592-874-1	Sequence 1, Appli
c	40	265	40.2	28804	3	US-09-096-942-2	Sequence 2, Appli
c	41	265	40.2	28804	3	US-09-096-867-2	Sequence 2, Appli
	42	265	40.2	4403765	4	US-09-103-840A-2	Sequence 2, Appli
	43	265	40.2	4411529	4	US-09-103-840A-1	Sequence 1, Appli
	44	264.5	40.1	3909	1	US-08-232-537-1	Sequence 1, Appli
	45	262.5	39.8	3999	4	US-09-351-224E-9	Sequence 9, Appli

ALIGNMENTS

```
RESULT 1
US-09-120-513-1
; Sequence 1, Application US/09120513
; Patent No. 6025160
;  GENERAL INFORMATION:
;    APPLICANT:  Brun, Kimberly
;    APPLICANT:  Chenery, Richard
;    APPLICANT:  Ellens, Harma
;    APPLICANT:  Field, John
;    APPLICANT:  Yue, Lin
;    TITLE OF INVENTION:  POLYNUCLEOTIDE AND POLYPEPTIDE
;    TITLE OF INVENTION:  SEQUENCES ENCODING RAT MDR1B2 AND
;    TITLE OF INVENTION:  SCREENING METHODS THEREOF
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road
;      CITY:  King of Prussia
;      STATE:  PA
;      COUNTRY:
;      ZIP:  19406
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/120,513
;      FILING DATE:  22-JUL-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  King, William T
;      REGISTRATION NUMBER:  30,954
;      REFERENCE/DOCKET NUMBER:  GP50008
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-5015
```

```
;      TELEFAX:  610-270-5090
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4233 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
US-09-120-513-1

Alignment Scores:
Pred. No.:                3.59e-55       Length:        4233
Score:                    472.00         Matches:       91
Percent Similarity:       91.06%         Conservative:  21
Best Local Similarity:    73.98%         Mismatches:    11
Query Match:              71.62%         Indels:        0
DB:                       3              Gaps:          0

US-09-873-409-3 (1-131) x US-09-120-513-1 (1-4233)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
             :::|||      ||||||||   :::|||||||||:::   |||:::   |||||||||
Db    1376 ATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTG 1435

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           ||||||||||||||||||||||||    |||||||||:::   |||||:::||||||||
Db    1436 GTGAGTCAGGAACCCGTGCTGTTTGCCACCACGATTGCCGAAAACATTCGCTATGGCCGA 1495

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
           ::::::||||||   :::|||:::|||:::|||   :::||||||||||||||||||||
Db    1496 GAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTC 1555

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           ||||||:::   |||:::|||||||||||||||||||||:::||||||||:::|||
Db    1556 ATCATGAAACTGCCCCACAAATTTAACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGT 1615

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1616 GGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAACCCCAAGATCCTT 1675

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
           :::|||||||||||||||||||||||||:::||||||::::::   |||||||||||
Db    1676 TTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTGGTTCAGGCCGCT 1735

Qy     122 LeuGluLys 124
           |||:::|||
Db    1736 CTGGATAAG 1744

RESULT 2
US-09-450-105-1
; Sequence 1, Application US/09450105
; Patent No. 6169166
; GENERAL INFORMATION:
;  APPLICANT: Kimberly Anne Brun
;  APPLICANT: Richard James Chenery
;  APPLICANT: Harma Ellens
;  APPLICANT: John Anthony Feild
;  APPLICANT: Lin Yue
;  TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
;  TITLE OF INVENTION: ENCODING RAT MDR1B2 AND SCREENING METHODS THEREOF
;  FILE REFERENCE: GP-50008-D1
;  CURRENT APPLICATION NUMBER: US/09/450,105
;  CURRENT FILING DATE: 1999-11-29
;  EARLIER APPLICATION NUMBER: 09/120,513
;  EARLIER FILING DATE: 1998-07-22
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 4233
;   TYPE: DNA
;   ORGANISM: HOMO SAPIENS
US-09-450-105-1

Alignment Scores:
Pred. No.:                3.59e-55       Length:        4233
Score:                    472.00         Matches:       91
Percent Similarity:       91.06%         Conservative:  21
Best Local Similarity:    73.98%         Mismatches:    11
Query Match:              71.62%         Indels:        0
DB:                       4              Gaps:          0

US-09-873-409-3 (1-131) x US-09-450-105-1 (1-4233)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
             :::|||      ||||||||   :::|||||||||:::   |||:::   |||||||||
Db    1376 ATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTG 1435

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           ||||||||||||||||||||||||    |||||||||:::   |||||:::||||||||
Db    1436 GTGAGTCAGGAACCCGTGCTGTTTGCCACCACGATTGCCGAAAACATTCGCTATGGCCGA 1495

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
           ::::::||||||   :::|||:::|||:::|||   :::||||||||||||||||||||
Db    1496 GAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTC 1555

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           ||||||:::   |||:::|||||||||||||||||||||:::||||||||:::|||
Db    1556 ATCATGAAACTGCCCCACAAATTTAACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGT 1615

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1616 GGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAACCCCAAGATCCTT 1675

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
           :::|||||||||||||||||||||||||:::||||||::::::   |||||||||||
Db    1676 TTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTGGTTCAGGCCGCT 1735

Qy     122 LeuGluLys 124
           |||:::|||
Db    1736 CTGGATAAG 1744

RESULT 3
US-08-784-649A-1
; Sequence 1, Application US/08784649A
; Patent No. 5830697
;  GENERAL INFORMATION:
;    APPLICANT:  Sikic, Branimir I
;    APPLICANT:  Chen, Gang
;    TITLE OF INVENTION:  P-GLYCOPROTEIN MUTANT RESISTANT TO
;    TITLE OF INVENTION:  CYCLOSPORIN MODULATION
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  2200 Sand Hill Road
;      CITY:  Menlo Park
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/784,649A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sherwood, Pamela J
;      REGISTRATION NUMBER:  Reg.No. 5830697 36,677
;      REFERENCE/DOCKET NUMBER:  06037/007001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-322-5070
;      TELEFAX:  415-854-0875
;  INFORMATION FOR SEQ ID NO:  1:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4264 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-784-649A-1

Alignment Scores:
Pred. No.:              1.21e-53      Length:        4264
Score:                  461.00        Matches:       90
Percent Similarity:     89.43%        Conservative:  20
Best Local Similarity:  73.17%        Mismatches:    13
Query Match:            69.95%        Indels:        0
DB:                     2             Gaps:          0

US-09-873-409-3 (1-131) x US-08-784-649A-1 (1-4264)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
          |||||       ||||||||   :::||||||||     |||:::   ||||||||
Db   1495 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 1554

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
          ||||||||||||||||||||||||   |||||||||:::   ||||||:::|||||||||
Db   1555 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 1614

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
          ::::::||||||   :::|||:::|||:::|||   :::|||||||||||||||||||||
Db   1615 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1674

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
          ||||||:::   |||:::||||||:::|||||||||||||||:::|||||||||:::|||
Db   1675 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 1734

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1735 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 1794

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
          :::||||||||||||||||||||||||||||:::||||||::::::     ||||||   |||
Db   1795 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 1854

Qy    122 LeuGluLys 124
          |||:::|||
Db   1855 CTGGATAAG 1863

RESULT 4
US-08-784-649A-5
; Sequence 5, Application US/08784649A
; Patent No. 5830697
;  GENERAL INFORMATION:
;    APPLICANT:  Sikic, Branimir I
;    APPLICANT:  Chen, Gang
;    TITLE OF INVENTION:  P-GLYCOPROTEIN MUTANT RESISTANT TO
;    TITLE OF INVENTION:  CYCLOSPORIN MODULATION
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  2200 Sand Hill Road
;      CITY:  Menlo Park
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/784,649A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sherwood, Pamela J
;      REGISTRATION NUMBER:  Reg.No. 5830697 36,677
;      REFERENCE/DOCKET NUMBER:  06037/007001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-322-5070
;      TELEFAX:  415-854-0875
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4264 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-784-649A-5

Alignment Scores:
Pred. No.:              1.21e-53      Length:        4264
Score:                  461.00        Matches:       90
Percent Similarity:     89.43%        Conservative:  20
Best Local Similarity:  73.17%        Mismatches:    13
Query Match:            69.95%        Indels:        0
DB:                     2             Gaps:          0

US-09-873-409-3 (1-131) x US-08-784-649A-5 (1-4264)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
          |||||       ||||||||   :::||||||||     |||:::   ||||||||
Db   1495 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 1554

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
          ||||||||||||||||||||||||   |||||||||:::   ||||||:::|||||||||
Db   1555 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 1614

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
          ::::::||||||   :::|||:::|||:::|||   :::|||||||||||||||||||||
Db   1615 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1674

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
          ||||||:::   |||:::||||||:::|||||||||||||||:::|||||||||:::|||
Db   1675 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 1734

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1735 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 1794

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
          :::||||||||||||||||||||||||||||:::||||||::::::     ||||||   |||
Db   1795 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 1854

Qy    122 LeuGluLys 124
          |||:::|||
Db   1855 CTGGATAAG 1863

RESULT 5
US-08-181-471-2
; Sequence 2, Application US/08181471
; Patent No. 5641508
;  GENERAL INFORMATION:
;    APPLICANT:  Li, Lingna
;    APPLICANT:  Lishko, Valeryi K.
;    TITLE OF INVENTION:  METHOD FOR DELIVERING BENEFICIAL
;    TITLE OF INVENTION:  COMPOSITIONS TO HAIR FOLLICLES
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Thomas Fitting
;      STREET:  12526 High Bluff Drive, Suite 300
;      CITY:  San Diego
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92130
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/181,471
;      FILING DATE:  13-JAN-1994
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/041,553
;      FILING DATE:  02-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitting, Thomas
;      REGISTRATION NUMBER:  34,163
;      REFERENCE/DOCKET NUMBER:  ANT0029P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-792-3680
;      TELEFAX:  619-792-8477
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4646 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  425..4267
US-08-181-471-2

Alignment Scores:
Pred. No.:              1.37e-53      Length:         4646
Score:                  461.00        Matches:        90
Percent Similarity:     89.43%        Conservative:   20
Best Local Similarity:  73.17%        Mismatches:     13
Query Match:            69.95%        Indels:         0
DB:                     1             Gaps:           0

US-09-873-409-3 (1-131) x US-08-181-471-2 (1-4646)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
             |||||      ||||||||   :::||||||||        |||:::   |||||||||
Db    1781 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 1840

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           ||||||||||||||||||||||||   |||||||||:::   ||||||:::|||||||||
Db    1841 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 1900

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
           ::::::||||||   :::|||:::|||:::|||   :::|||||||||||||||||||||
Db    1901 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1960

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           ||||||:::   |||:::||||||:::||||||||||||||||:::||||||||||:::|||
Db    1961 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 2020

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2021 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 2080

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
           :::|||||||||||||||||||||||||||:::||||||::::::   ||||||   |||
Db    2081 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 2140

Qy     122 LeuGluLys 124
           |||:::|||
Db    2141 CTGGATAAG 2149

RESULT 6
US-08-752-447-1
; Sequence 1, Application US/08752447
; Patent No. 5994088
;  GENERAL INFORMATION:
```

```
;    APPLICANT:  Mechetner, Eugene
;    APPLICANT:  Roninson, Igor B
;    TITLE OF INVENTION:  Methods and Reagents for Preparing and
;    TITLE OF INVENTION:  Using Immunoligcal Agents Specific for P-glycoprotein
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McDonnell Boehnen Hulbert & Berghoff Ltd.
;      STREET:  300 South Wacker Drive, Seventh Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/752,447
;      FILING DATE:  15-NOV-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5994088nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  95,1121
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-913-0001
;      TELEFAX:  312-913-9808
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4669 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  5'UTR
;      LOCATION:  1..424
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  425..4264
;    FEATURE:
;      NAME/KEY:  3'UTR
;      LOCATION:  4265..4669
US-08-752-447-1

Alignment Scores:
Pred. No.:              1.38e-53      Length:         4669
Score:                  461.00        Matches:        90
Percent Similarity:     89.43%        Conservative:   20
Best Local Similarity:  73.17%        Mismatches:     13
Query Match:            69.95%        Indels:         0
DB:                     2             Gaps:           0

US-09-873-409-3 (1-131) x US-08-752-447-1 (1-4669)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
             |||||      ||||||||   :::||||||||        |||:::   |||||||||
Db    1781 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 1840

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           ||||||||||||||||||||||||   |||||||||:::   ||||||:::|||||||||
Db    1841 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 1900

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
           ::::::||||||   :::|||:::|||:::|||   :::|||||||||||||||||||||
Db    1901 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1960

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           ||||||:::   |||:::||||||:::||||||||||||||||:::||||||||||:::|||
Db    1961 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 2020

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2021 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 2080

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
           :::||||||||||||||||||||||||:::||||||::::::    ||||||   |||
Db    2081 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 2140

Qy     122 LeuGluLys 124
           |||:::|||
Db    2141 CTGGATAAG 2149

RESULT 7
US-09-316-167-1
; Sequence 1, Application US/09316167
; Patent No. 6365357
;  GENERAL INFORMATION:
;    APPLICANT:  Mechetner, Eugene
;    APPLICANT:  Roninson, Igor B
;    TITLE OF INVENTION:  Methods and Reagents for Preparing and
;    TITLE OF INVENTION:  Using Immunoligcal Agents Specific for P-glycoprotein
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McDonnell Boehnen Hulbert & Berghoff Ltd.
;      STREET:  300 South Wacker Drive, Seventh Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/316,167
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/752,447
;      FILING DATE:  15-NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6365357nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  95,1121
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-913-0001
;      TELEFAX:  312-913-9808
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  4669 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  cDNA
;     FEATURE:
;       NAME/KEY:  5'UTR
;       LOCATION:  1..424
;     FEATURE:
;       NAME/KEY:  CDS
;       LOCATION:  425..4264
;     FEATURE:
;       NAME/KEY:  3'UTR
;       LOCATION:  4265..4669
US-09-316-167-1

Alignment Scores:
Pred. No.:              1.38e-53      Length:         4669
Score:                  461.00        Matches:        90
Percent Similarity:     89.43%        Conservative:   20
Best Local Similarity:  73.17%        Mismatches:     13
Query Match:            69.95%        Indels:         0
DB:                     4             Gaps:           0

US-09-873-409-3 (1-131) x US-09-316-167-1 (1-4669)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
            |||||      ||||||||    :::||||||||       |||:::   |||||||||
Db    1781 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 1840

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           ||||||||||||||||||||||||   ||||||||||:::   ||||||:::|||||||||
Db    1841 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 1900

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
           ::::::||||||   :::|||:::|||:::|||   :::||||||||||||||||||||||
Db    1901 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1960

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           ||||||:::   |||:::||||||:::||||||||||||||||:::|||||||||:::|||
Db    1961 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 2020

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2021 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 2080

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
           :::||||||||||||||||||||||||:::||||||::::::    ||||||   |||
Db    2081 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 2140

Qy     122 LeuGluLys 124
           |||:::|||
Db    2141 CTGGATAAG 2149

RESULT 8
5206352-3
;Patent No. 5206352
;    APPLICANT: Roninson, Igor B.;Pastan Ira H.;Gottesman,
;Michael M.
;    TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA
;SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
;    NUMBER OF SEQUENCES: 4
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/622,836
;      FILING DATE: 24-SEP-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 892,575
;      FILING DATE: 01-AUG-1986
;      APPLICATION NUMBER: 845,610
;      FILING DATE: 28-MAR-1986
;SEQ ID NO:3:
;      LENGTH: 4669
5206352-3

Alignment Scores:
Pred. No.:              1.38e-53      Length:         4669
Score:                  461.00        Matches:        90
Percent Similarity:     89.43%        Conservative:   20
Best Local Similarity:  73.17%        Mismatches:     13
Query Match:            69.95%        Indels:         0
DB:                     6             Gaps:           0

US-09-873-409-3 (1-131) x 5206352-3 (1-4669)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
            |||||      ||||||||    :::||||||||       |||:::   |||||||||
Db    1781 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 1840

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           ||||||||||||||||||||||||   ||||||||||:::   ||||||:::|||||||||
Db    1841 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 1900

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
           ::::::||||||   :::|||:::|||:::|||   :::||||||||||||||||||||||
Db    1901 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1960

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
```

```
                ||||||:::   |||:::||||||:::||||||||||||||:::|||||||||:::|||
Db     1961 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 2020

Qy       82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2021 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 2080

Qy      102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
            :::|||||||||||||||||||||||||:::||||||::::::   ||||||     |||
Db     2081 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 2140

Qy      122 LeuGluLys 124
            |||:::|||
Db     2141 CTGGATAAG 2149
```

RESULT 9
US-08-793-610-5
; Sequence 5, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
;   APPLICANT: BAUM, Christopher
;   APPLICANT: STOCKING-HARBERS, Carol
;   APPLICANT: OSTERTAG, Wolfram
;   TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEROF
;   TITLE OF INVENTION: FOR GENE TRANSFER
;   NUMBER OF SEQUENCES: 6
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
;   STREET: 655 Fifteenth Street N.W. Suite 330
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: U.S.A.
;   ZIP: 20005-5701
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/793,610
;   FILING DATE: 07-MAR-1997
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: DE P 44 31 973.8
;   FILING DATE: 08-SEP-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: DE 195 03 952.1
;   FILING DATE: 07-FEB-1995
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PCT/EP95/03175
;   FILING DATE: 10-AUG-1995
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Berman, Richard J.
;   REGISTRATION NUMBER: 39,105
;   REFERENCE/DOCKET NUMBER: P1614-7007
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (202)638-5000
;   TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 5:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 6505 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: circular
;   MOLECULE TYPE: DNA
US-08-793-610-5

```
Alignment Scores:
Pred. No.:              2.21e-53     Length:          6505
Score:                  461.00       Matches:         90
Percent Similarity:     89.43%       Conservative:    20
Best Local Similarity:  73.17%       Mismatches:      13
Query Match:            69.95%       Indels:          0
DB:                     2            Gaps:            0
```

US-09-873-409-3 (1-131) x US-08-793-610-5 (1-6505)

```
Qy        2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
            ||||||      |||||||||     :::|||||||||      |||:::   |||||||||
Db     3173 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 3232

Qy       22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
            ||||||||||||||||||||||||      |||||||||:::   ||||||:::|||||||||
Db     3233 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 3292

Qy       42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
            ::::::||||||   :::|||:::|||:::|||   :::|||||||||||||||||||||
Db     3293 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 3352

Qy       62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
            ||||||:::   |||:::||||||:::||||||||||||||:::|||||||||:::|||
Db     3353 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 3412

Qy       82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3413 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 3472

Qy      102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
            :::|||||||||||||||||||||||||:::||||||::::::   ||||||     |||
Db     3473 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 3532

Qy      122 LeuGluLys 124
            |||:::|||
Db     3533 CTGGATAAG 3541
```

RESULT 10
US-08-793-610-6
; Sequence 6, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
;   APPLICANT: BAUM, Christopher
;   APPLICANT: STOCKING-HARBERS, Carol
;   APPLICANT: OSTERTAG, Wolfram
;   TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEROF
;   TITLE OF INVENTION: FOR GENE TRANSFER
;   NUMBER OF SEQUENCES: 6
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
;   STREET: 655 Fifteenth Street N.W. Suite 330
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: U.S.A.
;   ZIP: 20005-5701
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/793,610
;   FILING DATE: 07-MAR-1997
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: DE P 44 31 973.8
;   FILING DATE: 08-SEP-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: DE 195 03 952.1
;   FILING DATE: 07-FEB-1995
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PCT/EP95/03175
;   FILING DATE: 10-AUG-1995
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Berman, Richard J.
;   REGISTRATION NUMBER: 39,105
;   REFERENCE/DOCKET NUMBER: P1614-7007
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (202)638-5000
;   TELEFAX: (202)638-4810

```
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9318 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA
US-08-793-610-6

Alignment Scores:
Pred. No.:              3.7e-53        Length:        9318
Score:                  461.00         Matches:       90
Percent Similarity:     89.43%         Conservative:  20
Best Local Similarity:  73.17%         Mismatches:    13
Query Match:            69.95%         Indels:        0
DB:                     2              Gaps:          0

US-09-873-409-3 (1-131) x US-08-793-610-6 (1-9318)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
           |||||      ||||||||   :::|||||||||      |||:::   |||||||||
Db    3132 GTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 3191

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           |||||||||||||||||||||||||||   ||||||||:::   ||||||:::|||||||||
Db    3192 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 3251

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
           ::::::||||||   :::|||:::|||:::|||   :::||||||||||||||||||||||
Db    3252 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 3311

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           ||||||:::   |||:::||||||:::||||||||||||||||:::||||||||||:::|||
Db    3312 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGT 3371

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3372 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 3431

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
           :::||||||||||||||||||||||||||:::||||||::::::   ||||||   |||
Db    3432 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 3491

Qy     122 LeuGluLys 124
           |||:::|||
Db    3492 CTGGATAAG 3500

RESULT 11
US-08-583-276-18
; Sequence 18, Application US/08583276
; Patent No. 5837536
;  GENERAL INFORMATION:
;    APPLICANT:  McDonagh, Kevin T.
;    APPLICANT:  Nienhuis, Arthur
;    APPLICANT:  Tolstoshev, Paul
;    TITLE OF INVENTION:  IMPROVED EXPRESSION OF HUMAN
;    TITLE OF INVENTION:  MULTIDRUG RESISTANCE GENES AND IMPROVED
;    TITLE OF INVENTION:  SELECTION OF CELLS TRANSDUCED WITH SUCH GENES
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carella, Byrne, Bain, Gilfillan,
;      ADDRESSEE:  Cecchi & Stewart
;      STREET:  6 Becker Farm Road
;      CITY:  Roseland
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch diskette
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  DW4.V2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/583,276
;      FILING DATE:  05-JAN-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/332,444
;      FILING DATE:  31-OCT-1994
;      APPLICATION NUMBER:  07/887,712
;      FILING DATE:  22-MAY-1992
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4669 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  singular
;      TOPOLOGY:  linear
;    MOLECULE TYPE:
;      DESCRIPTION:  Genomic DNA
US-08-583-276-18

Alignment Scores:
Pred. No.:              6.31e-52       Length:        4669
Score:                  449.00         Matches:       88
Percent Similarity:     87.80%         Conservative:  20
Best Local Similarity:  71.54%         Mismatches:    15
Query Match:            68.13%         Indels:        0
DB:                     2              Gaps:          0

US-09-873-409-3 (1-131) x US-08-583-276-18 (1-4669)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
           |||         ||||||||   :::|||||||||      |||:::   |||||||||
Db    1781 GTTCATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTG 1840

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           |||||||||||||||||||||||||   ||||||||:::   ||||||:::|||||||||
Db    1841 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGT 1900

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
           ::::::||||||   :::|||:::|||:::|||   :::||||||||||||||||||||||
Db    1901 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1960

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           ||||||:::   |||:::||||||:::||||||||||||||||:::|||   |||:::|||
Db    1961 ATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGGCCCCAGTTGAGT 2020

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2021 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 2080

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
           :::||||||||||||||||||||||||||:::||||||::::::   ||||||   |||
Db    2081 CTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCT 2140

Qy     122 LeuGluLys 124
           |||:::|||
Db    2141 CTGGATAAG 2149

RESULT 12
US-08-461-823-1
; Sequence 1, Application US/08461823
; Patent No. 5593840
;  GENERAL INFORMATION:
;    APPLICANT:  Bhatnagar, Satish K.
;    APPLICANT:  George Jr., Albert L.
;    APPLICANT:  Nazarenko, Irina
;    TITLE OF INVENTION:  AMPLIFICATION OF NUCLEIC ACID SEQUENCES
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OncorPharm, Inc.
;      STREET:  200 Perry Parkway
;      CITY:  Gaithersburg
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20877
```

```
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/461,823
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/168,621
;      FILING DATE:  16-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/010,433
;      FILING DATE:  27-JAN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Karta, Glenn E.
;      REGISTRATION NUMBER:  30,649
;      REFERENCE/DOCKET NUMBER:  PA-0012 CIP 2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301 527-2058
;      TELEFAX:  301 208-6997
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2726 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-461-823-1

Alignment Scores:
Pred. No.:              1.33e-41      Length:        2726
Score:                  372.00        Matches:       75
Percent Similarity:     77.78%        Conservative:  23
Best Local Similarity:  59.52%        Mismatches:    26
Query Match:            56.45%        Indels:        2
DB:                     1             Gaps:          1

US-09-873-409-3 (1-131) x US-08-461-823-1 (1-2726)

Qy        1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
            ::::::|||         :::|||:::   ||||||||||||   |||   |||:::|||
Db     1787 CTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGC 1846

Qy       21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
            :::|||||||||||||||:::||||||      :::|||:::   ||||||   ||||||
Db     1847 ATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGA 1906

Qy       41 ArgAspAsp------ValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla 58
               :::         |||:::   |||||||||:::   |||||||||:::|||||||||
Db     1907 GACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATA 1966

Qy       59 TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAla 78
            :::   ||||||         |||||||||::::::|||   ||||||:::||||||
Db     1967 CATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACT 2026

Qy       79 GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro 98
            |||:::||||||||||||||||||||||||||||||||||||||||||   |||
Db     2027 CAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCT 2086

Qy       99 LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 118
               ||||||:::|||||||||||||||||||||||||||:::||||||:::      |||
Db     2087 CATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTC 2146

Qy      119 GlnAlaAlaLeuGluLys 124
            |||   ||||||:::|||
Db     2147 CAAGAAGCCCTGGACAAA 2164

RESULT 13
US-08-996-545-1
; Sequence 1, Application US/08996545
; Patent No. 5928898
;  GENERAL INFORMATION:
;    APPLICANT:  Skatrud, Paul L.
;    APPLICANT:  de Waard, Maarten A.
;    APPLICANT:  Peery, Robert B.
;    APPLICANT:  Andrade, Alan C.
;    TITLE OF INVENTION:  Multiple Drug Resistance Gene atrD of
;    TITLE OF INVENTION:  Aspergillus Nidulans
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  U.S.
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/996,545
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Webster, Thomas D.
;      REGISTRATION NUMBER:  39,872
;      REFERENCE/DOCKET NUMBER:  X-11766
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-3334
;      TELEFAX:  317-276-2763
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4002 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..4002
US-08-996-545-1

Alignment Scores:
Pred. No.:              1.56e-40      Length:        4002
Score:                  366.00        Matches:       74
Percent Similarity:     73.60%        Conservative:  18
Best Local Similarity:  59.20%        Mismatches:    31
Query Match:            55.54%        Indels:        2
DB:                     2             Gaps:          1

US-09-873-409-3 (1-131) x US-08-996-545-1 (1-4002)

Qy        1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
            :::||||||      ||||||      ||||||:::      ||||||      :::
Db     3445 CTTGTTGATGGGAAGGACATAAGTAAACTAAATATCAACTCCTACCGCAGCTTTCTGTCA 3504

Qy       21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
            :::|||||||||||||||   |||:::      ||||||      ||||||      |||
Db     3505 CTGGTCAGCCAGGAGCCGACACTGTACCAGGGCACCATCAAGGAAAACATCTTACTTGGT 3564

Qy       41 ------ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla 58
                     |||||||||   :::|||   :::   :::|||   ::::::||||||
Db     3565 ATTGTCGAAGATGACGTACCGGAAGAATTCTTGATTAAGGCTTGCAAGGACGCTAATATC 3624

Qy       59 TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAla 78
            |||||||||||||||      |||      |||||||||:::||||||   ||||||
Db     3625 TACGACTTCATCATGTCGCTCCCGGAGGGCTTTAATACAGTTGTTGGCAGCAAGGGAGGC 3684
```

```
Qy      79 GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro 98
              :::||||||||||||||||||||:::|||||||||||||||||:::|||:::|||
Db    3685 ATGTTGTCTGGCGGCCAAAAGCAACGTGTGGCCATTGCCCGAGCCCTTCTTCGGGATCCC 3744

Qy      99 LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 118
           ||||||||:::||||||||||||||||||||||||||||||||||||:::      |||
Db    3745 AAAATCCTTCTTCTCGATGAAGCGACGTCAGCCCTCGACTCCGAGTCAGAAAAGGTCGTC 3804

Qy     119 GlnAlaAlaLeuGlu 123
           ||||||||||||:::
Db    3805 CAGGCGGCTTTGGAT 3819

RESULT 14
US-08-996-545-3
; Sequence 3, Application US/08996545
; Patent No. 5928898
;  GENERAL INFORMATION:
;    APPLICANT:  Skatrud, Paul L.
;    APPLICANT:  de Waard, Maarten A.
;    APPLICANT:  Peery, Robert B.
;    APPLICANT:  Andrade, Alan C.
;    TITLE OF INVENTION:  Multiple Drug Resistance Gene atrD of
;    TITLE OF INVENTION:  Aspergillus Nidulans
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  U.S.
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/996,545
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Webster, Thomas D.
;      REGISTRATION NUMBER:  39,872
;      REFERENCE/DOCKET NUMBER:  X-11766
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-3334
;      TELEFAX:  317-276-2763
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4002 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  mRNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-996-545-3

Alignment Scores:
Pred. No.:              1.56e-40      Length:        4002
Score:                  366.00        Matches:       74
Percent Similarity:     73.60%        Conservative:  18
Best Local Similarity:  59.20%        Mismatches:    31
Query Match:            55.54%        Indels:        2
DB:                     2             Gaps:          1

US-09-873-409-3 (1-131) x US-08-996-545-3 (1-4002)

Qy       1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
              :::||||||      ||||||        ||||||:::      ||||||      :::
Db    3445 CUUGUUGAUGGGAAGGACAUAAGUAAACUAAAUAUCAACUCCUACCGCAGCUUUCUGUCA 3504

Qy      21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
              :::||||||||||||||   |||:::      ||||||      ||||||      |||
Db    3505 CUGGUCAGCCAGGAGCCGACACUGUACCAGGGCACCAUCAAGGAAAACAUCUUACUUGGU 3564

Qy      41 ------ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla 58
                 |||||||||   :::|||   :::   :::|||   ::::::||||||
Db    3565 AUUGUCGAAGAUGACGUACCGGAAGAAUUCUUGAUUAAGGCUUGCAAGGACGCUAAUAUC 3624

Qy      59 TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAla 78
           ||||||||||||||      |||      |||||||||:::||||||      ||||||
Db    3625 UACGACUUCAUCAUGUCGCUCCCGGAGGGCUUUAAUACAGUUGUUGGCAGCAAGGGAGGC 3684

Qy      79 GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro 98
              :::|||||||||||||||||||:::|||||||||||||||||:::|||:::|||
Db    3685 AUGUUGUCUGGCGGCCAAAAGCAACGUGUGGCCAUUGCCCGAGCCCUUCUUCGGGAUCCC 3744

Qy      99 LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 118
           ||||||||:::||||||||||||||||||||||||||||||||||||:::      |||
Db    3745 AAAAUCCUUCUUCUCGAUGAAGCGACGUCAGCCCUCGACUCCGAGUCAGAAAAGGUCGUC 3804

Qy     119 GlnAlaAlaLeuGlu 123
           ||||||||||||:::
Db    3805 CAGGCGGCUUUGGAU 3819

RESULT 15
US-09-328-320-1
; Sequence 1, Application US/09328320
; Patent No. 6228615
;  GENERAL INFORMATION:
;    APPLICANT:  Skatrud, Paul L.
;    APPLICANT:  de Waard, Maarten A.
;    APPLICANT:  Peery, Robert B.
;    APPLICANT:  Andrade, Alan C.
;    TITLE OF INVENTION:  Multiple Drug Resistance Gene atrD of
;    TITLE OF INVENTION:  Aspergillus Nidulans
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  U.S.
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/328,320
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/996,545
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Webster, Thomas D.
;      REGISTRATION NUMBER:  39,872
;      REFERENCE/DOCKET NUMBER:  X-11766
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-3334
;      TELEFAX:  317-276-2763
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4002 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..4002
```

```
US-09-328-320-1

Alignment Scores:
Pred. No.:              1.56e-40      Length:         4002
Score:                  366.00        Matches:        74
Percent Similarity:     73.60%        Conservative:   18
Best Local Similarity:  59.20%        Mismatches:     31
Query Match:            55.54%        Indels:         2
DB:                     4             Gaps:           1

US-09-873-409-3 (1-131) x US-09-328-320-1 (1-4002)

Qy       1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
           :::||||||      ||||||      ||||||:::      ||||||      :::
Db    3445 CTTGTTGATGGGAAGGACATAAGTAAACTAAATATCAACTCCTACCGCAGCTTTCTGTCA 3504

Qy      21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
           :::||||||||||||||||||   |||:::      ||||||   ||||||      |||
Db    3505 CTGGTCAGCCAGGAGCCGACACTGTACCAGGGCACCATCAAGGAAAACATCTTACTTGGT 3564

Qy      41 ------ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla 58
                 |||||||||   :::|||   :::   :::|||   ::::::||||||
Db    3565 ATTGTCGAAGATGACGTACCGGAAGAATTCTTGATTAAGGCTTGCAAGGACGCTAATATC 3624

Qy      59 TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAla 78
           |||||||||||||||      |||      |||||||||:::||||||   ||||||
Db    3625 TACGACTTCATCATGTCGCTCCCGGAGGGCTTTAATACAGTTGTTGGCAGCAAGGGAGGC 3684

Qy      79 GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro 98
              :::||||||||||||||||||||||:::||||||||||||||||||:::|||:::|||
Db    3685 ATGTTGTCTGGCGGCCAAAAGCAACGTGTGGCCATTGCCCGAGCCCTTCTTCGGGATCCC 3744

Qy      99 LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 118
           |||||||||:::|||||||||||||||||||||||||||||||||:::         |||
Db    3745 AAAATCCTTCTTCTCGATGAAGCGACGTCAGCCCTCGACTCCGAGTCAGAAAAGGTCGTC 3804

Qy     119 GlnAlaAlaLeuGlu 123
           ||||||||||||:::
Db    3805 CAGGCGGCTTTGGAT 3819

Search completed: March 31, 2003, 15:23:26
Job time : 39.1843 secs
```

```
                    GenCore version 5.1.4_p5_4578
           Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 30, 2003, 03:22:44 ; Search time 901.042 Seconds
                                           (without alignments)
                                           2354.618 Million cell updates/sec

Title:          US-09-873-409-3
Perfect score:  659
Sequence:       1 MVDENDIRALNVRHYRDHIG..........SESKSAVQAALEKDTPRYSF 131

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters:      32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115420_19240/app_query.fasta_1.7544
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_30544_@runat_27032003_115420_19240 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      EST:*
                1:  em_estba:*
                2:  em_esthum:*
                3:  em_estin:*
                4:  em_estmu:*
                5:  em_estov:*
                6:  em_estpl:*
                7:  em_estro:*
                8:  em_htc:*
                9:  gb_est1:*
                10:  gb_est2:*
                11:  gb_htc:*
                12:  gb_est3:*
                13:  gb_est4:*
                14:  gb_est5:*
                15:  em_estfun:*
                16:  em_estom:*
                17:  gb_gss:*
                18:  em_gss_hum:*
                19:  em_gss_inv:*
                20:  em_gss_pln:*
                21:  em_gss_vrt:*
                22:  em_gss_fun:*
                23:  em_gss_mam:*
                24:  em_gss_mus:*
                25:  em_gss_other:*
                26:  em_gss_pro:*
                27:  em_gss_rod:*
```

```
     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                                SUMMARIES
                     %
Result            Query
  No.    Score   Match Length DB  ID                    Description
-----------------------------------------------------------------------------
     1    618     93.8    760  13  BM016204              BM016204 603642659
     2    578     87.7    559  12  BF692596              BF692596 602248949
     3    524     79.5    545   9  AL040762              AL040762 DKFZp434C
     4    490     74.4    669  13  BM013981              BM013981 603639619
     5    464.5   70.5   1033  14  BM926413              BM926413 AGENCOURT
     6    445     67.5    550  13  BJ075336              BJ075336 BJ075336
c    7    428     64.9   1098  17  CNS051EN              AL316760 Tetraodon
     8    393     59.6    550  13  BJ388678              BJ388678 BJ388678
     9    379     57.5    520  13  BJ336809              BJ336809 BJ336809
    10    378.5   57.4    588   9  AI666987              AI666987 fc24c03.y
    11    378.5   57.4    589   9  AI666988              AI666988 fc24c04.y
    12    377     57.2    542  13  BJ339421              BJ339421 BJ339421
    13    377     57.2    547  13  BJ333166              BJ333166 BJ333166
c   14    374     56.8    589  13  BJ356968              BJ356968 BJ356968
c   15    374     56.8    608  13  BJ356648              BJ356648 BJ356648
c   16    374     56.8    656  13  BJ401771              BJ401771 BJ401771
c   17    374     56.8    659  13  BJ350425              BJ350425 BJ350425
c   18    374     56.8    687  13  BJ375176              BJ375176 BJ375176
c   19    374     56.8    688  13  BJ405288              BJ405288 BJ405288
c   20    374     56.8    690  13  BJ408755              BJ408755 BJ408755
c   21    374     56.8    692  13  BJ407928              BJ407928 BJ407928
c   22    374     56.8    696  13  BJ348083              BJ348083 BJ348083
c   23    374     56.8    722  13  BJ406558              BJ406558 BJ406558
c   24    374     56.8    731  13  BJ354552              BJ354552 BJ354552
c   25    374     56.8    733  13  BJ376217              BJ376217 BJ376217
c   26    374     56.8    739  13  BJ404475              BJ404475 BJ404475
c   27    374     56.8    741  13  BJ347139              BJ347139 BJ347139
c   28    374     56.8    742  13  BJ372910              BJ372910 BJ372910
c   29    374     56.8    743  13  BJ408863              BJ408863 BJ408863
c   30    374     56.8    745  13  BJ405178              BJ405178 BJ405178
c   31    374     56.8    758  13  BJ357956              BJ357956 BJ357956
c   32    374     56.8    783  13  BJ352760              BJ352760 BJ352760
c   33    374     56.8    784  13  BJ405287              BJ405287 BJ405287
c   34    374     56.8    792  13  BJ352404              BJ352404 BJ352404
c   35    374     56.8    814  13  BJ352304              BJ352304 BJ352304
c   36    374     56.8    820  13  BJ356540              BJ356540 BJ356540
c   37    374     56.8    835  13  BJ349604              BJ349604 BJ349604
c   38    374     56.8    838  13  BJ355307              BJ355307 BJ355307
    39    372     56.4    943   9  AL520322              AL520322 AL520322
    40    370     56.1    681  14  BU006064              BU006064 QGG9N16.y
    41    370     56.1    715  14  BQ869512              BQ869512 QGD6G09.y
    42    370     56.1    894  12  BF584668              BF584668 602098406
    43    369     56.0    508  14  BQ818890              BQ818890 1030074A0
    44    369     56.0    635  13  BI716519              BI716519 1031010G1
c   45    368.5   55.9    871   9  AL582625              AL582625 AL582625

                                ALIGNMENTS

RESULT 1
BM016204
LOCUS       BM016204                 760 bp    mRNA    linear   EST 30-OCT-2001
DEFINITION  603642659F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5418615 5',
            mRNA sequence.
ACCESSION   BM016204
VERSION     BM016204.1  GI:16530558
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 760)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
```

```
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12068  row: e  column: 16
            High quality sequence stop: 740.
FEATURES             Location/Qualifiers
     source          1. .760
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5418615"
                     /clone_lib="NIH_MGC_87"
                     /tissue_type="mammary adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.383 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      229 a    150 c    187 g    193 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:               5.63e-73      Length:        760
Score:                   618.00        Matches:       128
Percent Similarity:      96.97%        Conservative:  0
Best Local Similarity:   96.97%        Mismatches:    2
Query Match:             93.78%        Indels:        2
DB:                      13            Gaps:          0

US-09-873-409-3 (1-131) x BM016204 (1-760)

Qy       1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||    |||
Db     246 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATNGGA 305

Qy      21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     306 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 365

Qy      41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     366 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 425

Qy      61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     426 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 485

Qy      81 SerGlyGlyGlnLysGlnArgIleAlaIle-AlaArgAlaLeuValArgAsnProLysIl 100
           |||||||||||||||||||||||||||||| |||||||||||||||||||||||||||||
Db     486 AGTGGAGGGCAGAAACAGAGGATCGCAATTGGCTCGTGCCTTAGTTCGAAACCCCAAGAT 545

Qy     100 eLeuIleLeu-AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnA 120
           ||||||     ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     546 TCTGATTTTCAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAG 605

Qy     120 laAlaLeuGluLysAspThrProArgTyrSer 130
           ||||||||||||||||||||||||||||||||
Db     606 CTGCACTGGAGAAGGATACCCCCAGGTATTCC 637

RESULT 2
BF692596
LOCUS       BF692596                 559 bp    mRNA    linear   EST 22-DEC-2000
DEFINITION  602248949F1 NIH_MGC_62 Homo sapiens cDNA clone IMAGE:4334100 5',
            mRNA sequence.
ACCESSION   BF692596
VERSION     BF692596.1  GI:11978004
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 559)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1203  row: a  column: 13
            High quality sequence stop: 555.
FEATURES             Location/Qualifiers
     source          1. .559
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4334100"
                     /clone_lib="NIH_MGC_62"
                     /tissue_type="melanotic melanoma, high MDR"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: skin; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc);
                     Double-stranded cDNA was prepared from cell line RNA.  5'
                     and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).   Average
                     insert size 1.75 kb (range 0.9-4.0 kb).  15/15 colonies
                     contained  inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA)."
BASE COUNT      161 a    118 c    141 g    139 t
ORIGIN

Alignment Scores:
Pred. No.:               9.32e-68      Length:        559
Score:                   578.00        Matches:       126
Percent Similarity:      96.95%        Conservative:  1
Best Local Similarity:   96.18%        Mismatches:    3
Query Match:             87.71%        Indels:        3
DB:                      12            Gaps:          0

US-09-873-409-3 (1-131) x BF692596 (1-559)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
           |||||||||||||||||||||||||||||||||||||||||||||||||||    ||||||
Db      28 GTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATAT-GGAGTG 86

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           ||||||||||||||     |||||||||||||||||||||||||||||||||||||||||
Db      87 GTTAGTCAAGAGCCTGT-TTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGA 145

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     146 GATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCATATGATTTT 205

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLys-GlyAlaGlnMetSe 81
           ||||||||||||||||||||||||||||||||||||||||||||     ||||||||||||
Db     206 ATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAACAGGAGCTCAAATGAG 265

Qy      81 rGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLe 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     266 TGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATTCT 325

Qy     101 uIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAl 121
```

```
              ||||||||||||||||||||||||||||||||||||||||:::||||||||||||||||
Db      326 GATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCGAGTCAGCTGTTCAAGCTGC 385

Qy      121 aLeuGluLysAspThrProArgTyrSerPhe 131
              |||||||||||||||||||||||||||||||
Db      386 ACTGGAGAAGGATACCCCCAGGTATTCATTT 416

RESULT 3
AL040762
LOCUS       AL040762                 545 bp    mRNA    linear   EST 29-FEB-2000
DEFINITION  DKFZp434C1815_r1 434 (synonym: htes3) Homo sapiens cDNA clone
            DKFZp434C1815 5', mRNA sequence.
ACCESSION   AL040762
VERSION     AL040762.1  GI:5409708
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 545)
  AUTHORS   Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
            Wiemann,S.
  TITLE     EST (Bloecker, et al.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Bloecker H
            MIPS
            Am Klopferspitz 18a D-82152 Martinsried, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by GBF (National Research Centre for Biotechnology Ltd.,
            Braunschweig/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            s1 sequence also available.
            This clone (DKFZp434C1815) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1..545
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp434C1815"
                     /clone_lib="434 (synonym: htes3)"
                     /tissue_type="testis"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT      171 a     89 c    138 g    147 t
ORIGIN

Alignment Scores:
Pred. No.:             1.76e-60      Length:          545
Score:                 524.00        Matches:         103
Percent Similarity:    100.00%       Conservative:    0
Best Local Similarity: 100.00%       Mismatches:      0
Query Match:           79.51%        Indels:          0
DB:                    9             Gaps:            0

US-09-873-409-3 (1-131) x AL040762 (1-545)

Qy        1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      237 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 296

Qy       21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      297 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 356

Qy       41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      357 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 416

Qy       61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
```

```
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      417 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 476

Qy       81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      477 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 536

Qy      101 LeuIleLeu 103
            |||||||||
Db      537 CTGATTTTA 545

RESULT 4
BM013981
LOCUS       BM013981                 669 bp    mRNA    linear   EST 30-OCT-2001
DEFINITION  603639619F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5415556 5',
            mRNA sequence.
ACCESSION   BM013981
VERSION     BM013981.1  GI:16528335
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 669)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12060  row: f  column: 05
            High quality sequence stop: 664.
FEATURES             Location/Qualifiers
     source          1..669
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5415556"
                     /clone_lib="NIH_MGC_87"
                     /tissue_type="mammary adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.383 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      203 a    122 c    162 g    182 t
ORIGIN

Alignment Scores:
Pred. No.:             9.06e-56      Length:          669
Score:                 490.00        Matches:         98
Percent Similarity:    98.99%        Conservative:    0
Best Local Similarity: 98.99%        Mismatches:      0
Query Match:           74.36%        Indels:          1
DB:                    13            Gaps:            0

US-09-873-409-3 (1-131) x BM013981 (1-669)

Qy        1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      374 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 433

Qy       21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      434 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 493

Qy       41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     494 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 553

Qy      61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     554 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 613

Qy      81 SerGlyGlyGlnLysGlnArg-IleAlaIleAlaArgAlaLeuValArgAsnPro 98
           |||||||||||||||||||| ||||||||||||||||||||||||||||||||||
Db     614 AGTGGAGGGCAGAAACAGAGGCATCGCAATTGCTCGTGCCTTAGTTCGAAACCCC 668
```

```
RESULT 5
BM926413
LOCUS       BM926413               1033 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6600788 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5764845
            5', mRNA sequence.
ACCESSION   BM926413
VERSION     BM926413.1  GI:19376792
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1033)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12818  row: o  column: 22
            High quality sequence start: 3
            High quality sequence stop: 686.
FEATURES             Location/Qualifiers
     source          1..1033
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5764845"
                     /clone_lib="NIH_MGC_114"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
                     male_brains, age range 23-27 yo. Library is oligo-dT
                     primed and directionally cloned (EcoRV site is  destroyed
                     upon cloning).  Average insert size 1.5 kb, insert size
                     range 1-3 kb.  Library is normalized and enriched for
                     full-length  clones and was constructed by C. Gruber
                     (Invitrogen).  Research  Genetics tracking code 019. Note:
                     this is a NIH_MGC Library."
BASE COUNT      289 a    212 c    218 g    312 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:               4.54e-52      Length:        1033
Score:                   464.50        Matches:       111
Percent Similarity:      87.02%        Conservative:  3
Best Local Similarity:   84.73%        Mismatches:    17
Query Match:             70.49%        Indels:        4
DB:                      14            Gaps:          0

US-09-873-409-3 (1-131) x BM926413 (1-1033)

Qy       1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     534 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 593

Qy      21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
```

```
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     594 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 653

Qy      41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
           ||||||.|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     654 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 713

Qy      61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
           ||||||||||||||||||||| ||||||||||||||||||||||   ||||||   |||
Db     714 TTTATCATGGAGTTTCCTAATAA-TTTAATACATTGGTAGGGGAANAAGGAGCTCCAATG 772

Qy      81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
           ||||||   |||      ||||||  ||||||  |||||||||||||||||||:::
Db     773 AGTGGAAGCCAGAACCG-AGGATCCCAATTGCTCCTGCCTTAGTTCGAAACCCCAGGAT- 830

Qy     101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
           |||   ||||||||||||||||||||||||||||||||:::          |||   |||
Db     831 CTGAATTTAGATGAAGCTACGTCTGCCCTGGATTCAGAAACCAGTCAACTGGTCA--GCT 888

Qy     121 AlaLeuGluLysAspThrProArgTyrSerPhe 131
              |||||||||:::||||||||       |||
Db     889 GGCCTGGAGAAGGAAACCCCCCGGNAATTCTTT 921
```

```
RESULT 6
BJ075336
LOCUS       BJ075336                550 bp    mRNA    linear   EST 11-DEC-2001
DEFINITION  BJ075336 NIBB Mochii normalized Xenopus tailbud library Xenopus
            laevis cDNA clone XL057d10 5', mRNA sequence.
ACCESSION   BJ075336
VERSION     BJ075336.1  GI:17520252
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 550)
  AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
            ,Y.
  TITLE     Expressed genes in X. laevis embryo
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1..550
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="XL057d10"
                     /clone_lib="NIBB Mochii normalized Xenopus tailbud
                     library"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 25"
                     /note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
                     were oligo-dT primed and directionally cloned. Staging
                     according to  Nieuwkoop and Faber. Library is subtracted
                     and was constructed by N. Garrett  and A.M. Zorn,
                     (Wellcome/CRC Institute). "
BASE COUNT      162 a    109 c    144 g    135 t
ORIGIN

Alignment Scores:
Pred. No.:               8.26e-50      Length:        550
Score:                   445.00        Matches:       86
Percent Similarity:      90.00%        Conservative:  22
Best Local Similarity:   71.67%        Mismatches:    12
Query Match:             67.53%        Indels:        0
DB:                      13            Gaps:          0
```

```
US-09-873-409-3 (1-131) x BJ075336 (1-550)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
          :::|||   :::||||||||:::|||||||||:::         |||::::::||||||||
Db    191 TTGGATGGACACGATATCCGTTCACTAAATGTAAAATGGTTGAGGGAAAATATTGGCGTA 250

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
          |||||||||||||||||||||||||||||||||||||:::   ||||||:::||||||
Db    251 GTGAGTCAGGAGCCAGTGCTCTTTGGCACAACCATTGCAGAAAATATCCGCTACGGAGGG 310

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
          :::   |||||||||   |||:::|||:::||||||:::||||||||||||:::||||||
Db    311 GAATCAGTAACAGACAACGAAATTGAACAAGCGGCAAAAGAAGCCAATGCCTTTGATTTT 370

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
          |||            |||:::||||||   |||:::|||||||||:::|||||||||:::|||
Db    371 ATATCCAGGCTTCCTGATAAATTTAAGACGATGGTGGGAGAAAGAGGTGCGCAGTTGAGT 430

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    431 GGGGGTCAGAAGCAGCGGATAGCCATTGCACGCGCCCTGGTACGGAACCCTAAGATTCTA 490

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
          :::||||||||||||||||||||||||||||:::::::|||::::::   |||||||||||
Db    491 CTTCTTGATGAAGCCACATCAGCTCTGGATACACAAAGTGAAGCTATTGTACAAGCTGCC 550
```

```
RESULT 7
CNS051EN/c
LOCUS       CNS051EN                1098 bp    DNA     linear   GSS 26-JUL-2000
DEFINITION  Tetraodon nigroviridis genome survey sequence T7 end of clone
            052D12 of library A from Tetraodon nigroviridis, genomic survey
            sequence.
ACCESSION   AL316760
VERSION     AL316760.1  GI:9549644
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis.
  ORGANISM  Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodon.
REFERENCE   1  (bases 1 to 1098)
  AUTHORS   Roest Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
  TITLE     Estimate of human gene number provided by genome-wide analysis
            using Tetraodon nigroviridis DNA sequence
  JOURNAL   Nat. Genet. 25 (2), 235-238 (2000)
  MEDLINE   20296633
   PUBMED   10835645
REFERENCE   2  (bases 1 to 1098)
  AUTHORS   Crollius,H.R., Jaillon,O., Dasilva,C., Ozouf-Costaz,C., Fizames,C.,
            Fischer,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W.,
            Bernot,A. and Weissenbach,J.
  TITLE     Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
  JOURNAL   Genome Res. 10 (7), 939-949 (2000)
  MEDLINE   20359837
   PUBMED   10899143
REFERENCE   3  (bases 1 to 1098)
  AUTHORS   Genoscope.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-APR-2000)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/Tetraodon.
FEATURES             Location/Qualifiers
     source          1..1098
                     /organism="Tetraodon nigroviridis"
                     /db_xref="taxon:99883"
                     /clone="052D12"
                     /clone_lib="A"
                     /note="Genoscope sequence ID : C0AA052DB06C1~end : T7"
BASE COUNT      232 a    281 c    312 g    260 t     13 others
ORIGIN

Alignment Scores:
Pred. No.:              4.19e-47      Length:          1098
Score:                  428.00        Matches:         89
Percent Similarity:     75.34%        Conservative:    21
Best Local Similarity:  60.96%        Mismatches:      13
Query Match:            64.95%        Indels:          23
DB:                     17            Gaps:            1

US-09-873-409-3 (1-131) x CNS051EN (1-1098)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
          :::|||   :::|||||||||:::||||||:::|||:::    |||:::   |||||||||
Db   1005 ATCGACGGTCACGACATCCGCTCTCTCAACATCCGCTACCTGAGAGAGATGATCGGGGTG 946

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
          |||||||||||||||:::||||||   |||||||||:::   ||||||:::|||||||||
Db    945 GTGAGTCAGGAGCCCATCCTCTTCGCCACCACCATCGCCGAGAACATCCGATACGGCCGA 886

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
             |||||||||   ||||||:::|||||||||   :::|||:::|||||||||||||||
Db    885 CTGGACGTGACGCAGGAGGAGATCGAACGAGCCACCAAAGAGTCCAACGCTTACGACTTC 826

Qy     62 IleMetGluPheProAsn------------------------------------------ 67
          ||||||      |||:::
Db    825 ATCATGAGCCTTCCAGATGTGAGAGCCGGTTGGTCCCCGCCTCGGTCCCACCCTGGCTGA 766

Qy     68 ------------------------LysPheAsnThrLeuValGlyGluLysGlyAlaGl 79
                                  ||||||   |||||||||||::::::|||   ||
Db    765 TCTCTGCTCCGTCGTCTCTGAGCAGAAGTTTGAGACGCTGGTGGGAGACCGCGGGACTCA 706

Qy     79 nMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLy 99
          |:::||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    705 GCTGAGCGGAGGACAGAAGCAGAGGATCGCCATCGCCCGAGCTCTGGTCCGCAACCCCAA 646

Qy     99 sIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGl 119
          ||||||:::||||||||||||||||||||||||||||:::||||||::::::   |||||
Db    645 AATCCTGCTGCTGGATGAAGCCACGTCCGCCCTGGATGCTGAGAGCGAGACCATCGTACA 586

Qy    119 nAlaAlaLeuGluLys 124
          ||||||||||:::|||
Db    585 GGCTGCTCTGGACAAG 570
```

```
RESULT 8
BJ388678
LOCUS       BJ388678                 550 bp    mRNA    linear   EST 08-MAR-2002
DEFINITION  BJ388678 Dictyostelium discoideum cDNA library, SF Dictyostelium
            discoideum cDNA clone dds6p15 5', mRNA sequence.
ACCESSION   BJ388678
VERSION     BJ388678.1  GI:19299764
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 550)
  AUTHORS   Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
  TITLE     Full length cDNA of Dictyostelium discoideum at the slug stage
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1..550
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
```

```
                 /clone="dds6p15"
                 /clone_lib="Dictyostelium discoideum cDNA library, SF"
                 /sex="mat A"
                 /dev_stage="Slug stage"
BASE COUNT      190 a     77 c    113 g    170 t
ORIGIN

Alignment Scores:
Pred. No.:               8.66e-43        Length:         550
Score:                   393.00          Matches:        76
Percent Similarity:      78.86%          Conservative:   21
Best Local Similarity:   61.79%          Mismatches:     26
Query Match:             59.64%          Indels:         0
DB:                      13              Gaps:           0

US-09-873-409-3 (1-131) x BJ388678 (1-550)

Qy        1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
              ::::::|||      |||||||||      ||||||||      |||      ||||||
Db       96 TTATTGGATGGAGAAGATATTAGAAAATTCAATGTCAGAGGGTTACGTCAAAAAATTGGT 155

Qy       21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
              :::|||:::||||||||||||||||||   |||||||||||   ||||||:::||||||
Db      156 TTAGTCAATCAAGAGCCTGTATTATTCGCCACAACCATTAGTGAAAATATTCGTTATGGT 215

Qy       41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
              ::::::       |||   :::|||:::|||   ||||||:::   |||||||||:::
Db      216 AAAGAAGGTGCCACTCAAGATGAAATTGAAGAAGCAGCTAAATTAGCAAATGCCCATAGT 275

Qy       61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
              ||||||   :::   |||      :::||||||||||||||||||||||||   ||||||
Db      276 TTTATATCACAATTACCACAGGGTTATAATACTTTGGTAGGTGAGAAAGGTGTCCAAATG 335

Qy       81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
              ||||||||||||:::|||||||||||||||||||||||:::::::::||||||   |||
Db      336 TCTGGTGGTCAAAGACAACGTATAGCAATTGCTCGTGCCGTTATTAAAAATCCAAATATA 395

Qy      101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
              |||:::|||||||||:::||||||||||||||||:::||||||      ||||||
Db      396 TTACTCTTGGATGAATCAACATCTGCATTGGATGCAGAGTCTACAAAATTAGTACAAGAG 455

Qy      121 AlaLeuGlu 123
              ||||||:::
Db      456 GCATTAGAT 464

RESULT 9
BJ336809
LOCUS       BJ336809                 520 bp    mRNA    linear   EST 05-MAR-2002
DEFINITION  BJ336809 Dictyostelium discoideum cDNA library, AF Dictyostelium
            discoideum cDNA clone dda55p03 5', mRNA sequence.
ACCESSION   BJ336809
VERSION     BJ336809.1  GI:19166939
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 520)
  AUTHORS   Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
  TITLE     Full length cDNA of Dictyostelium discoideum at the aggregation
            stage
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1..520
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="dda55p03"
                     /clone_lib="Dictyostelium discoideum cDNA library, AF"
                     /sex="mat A"
                     /dev_stage="Aggregation stage"
BASE COUNT      178 a     78 c    106 g    158 t
ORIGIN

Alignment Scores:
Pred. No.:               6.23e-41        Length:         520
Score:                   379.00          Matches:        73
Percent Similarity:      78.81%          Conservative:   20
Best Local Similarity:   61.86%          Mismatches:     25
Query Match:             57.51%          Indels:         0
DB:                      13              Gaps:           0

US-09-873-409-3 (1-131) x BJ336809 (1-520)

Qy        1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
              ::::::|||      |||||||||      ||||||||      |||      ||||||
Db      165 TTATTGGATGGAGAAGATATTAGAAAATTCAATGTCAGAGGGTTACGTCAAAAAATTGGT 224

Qy       21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
              :::|||:::||||||||||||||||||   |||||||||||   ||||||:::||||||
Db      225 TTAGTCAATCAAGAGCCTGTATTATTCGCCACAACCATTAGTGAAAATATTCGTTATGGT 284

Qy       41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
              ::::::       |||   :::|||:::|||   ||||||:::   |||||||||:::
Db      285 AAAGAAGGTGCCACTCAAGATGAAATTGAAGAAGCAGCTAAATTAGCAAATGCCCATAGT 344

Qy       61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
              ||||||   :::   |||      :::||||||||||||||||||||||||   ||||||
Db      345 TTTATATCACAATTACCACAGGGTTATAATACTTTGGTAGGTGAGAAAGGTGTCCAAATG 404

Qy       81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
              ||||||||||||:::|||||||||||||||||||||||:::::::::||||||   |||
Db      405 TCTGGTGGTCAAAGACAACGTATAGCAATTGCTCGTGCCGTTATTAAAAATCCAAATATA 464

Qy      101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 118
              |||:::|||||||||:::||||||||||||||||:::||||||      |||
Db      465 TTACTCTTGGATGAATCAACATCTGCATTGGATGCAGAGTCTACAAAATTAGTA 518

RESULT 10
AI666987
LOCUS       AI666987                 588 bp    mRNA    linear   EST 07-JUN-2001
DEFINITION  fc24c03.y1 Zebrafish WashU MPIMG EST Danio rerio cDNA clone
            IMAGE:3722308 5' similar to TR:Q13040 Q13040 ATP-BINDING CASSETTE
            PROTEIN ;, mRNA sequence.
ACCESSION   AI666987
VERSION     AI666987.1  GI:4805343
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 588)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
            ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
            ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
            Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
            Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
            and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1998
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            cDNA Library Preparation: Matthew Clark. cDNA Library Arrayed by:
            Matthew Clark. DNA Sequencing by: Washington University Genome
            Sequencing Center Clone distribution: Genome Systems, St. Louis,
```

```
         Missouri (web address: www.genomesystems.com) (email contact:
         info@genomesystems.com) and Research Genetics, Huntsville, Alabama
         (web address: www.resgen.com) (email contact: info@resgen.com) and
         RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address:
         www.rzpd.de)
         Seq primer: T3 ET from Amersham
         High quality sequence stop: 481.
FEATURES             Location/Qualifiers
     source          1. .588
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:3722308"
                     /clone_lib="Zebrafish WashU MPIMG EST"
                     /sex="mixed"
                     /tissue_type="26 somite embryos, adult livers, shield
                     stage embryos"
                     /lab_host="XL1-blue MRF"
                     /note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; 1st
                     strand cDNA was primed with a Not I - oligo(dT)15 primer
                     [5'pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT3'];
                     double-stranded cDNA was ligated to Sal I adaptors (BRL),
                     digested with Not I and cloned into the Not I and Sal I
                     sites of the pSPORT1 vector (BRL). Library was constructed
                     by Matthew Clark (Lehrach lab; ICRF, London and Max Planck
                     Institut fuer Molekulare Genetik,Berlin). cDNAs for EST
                     analysis were selected following oligonucleotide
                     hybridization fingerprinting of arrayed clones from
                     zebrafish late somitogenesis (26 ss), adult liver or
                     embryonic shield stage (5.6 h) libraries.  Fingerprint
                     data were used to computationally cluster cDNAs, and a
                     single cDNA from each cluster was chosen for sequencing.
                     In some cases multiple members of the same cluster were
                     sequenced to assess clustering parameters or single clones
                     were sequenced additional times to assess quality
                     control."
BASE COUNT      149 a    138 c    167 g    132 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              8.61e-41      Length:         588
Score:                  378.50        Matches:        79
Percent Similarity:     77.78%        Conservative:   19
Best Local Similarity:  62.70%        Mismatches:     25
Query Match:            57.44%        Indels:         3
DB:                     9             Gaps:           1

US-09-873-409-3 (1-131) x AI666987 (1-588)

Qy        2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
            |||||    :::|||||||| ||||||            |||  |||||||||
Db        1 GTTGATGGTCATGATATTCGAGACTTGAATCCTTACTGGTTAAGGAGTCACATTGGTACT 60

Qy       22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
            ||||||||||||||||||||||||       :::|||:::   ||||||   ||||||
Db       61 GTAAGCCAGGAACCGGTCCTTTTCTCCTGTTCAATTGCTGAGAACATTGCTTATGGAGCG 120

Qy       42 AspAsp---------ValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla 58
               |||          ||||||  |||::::::  |||||||||:::  |||||||||
Db      121 CCGGATCCGAGTCAGGTGACAGTGGAGGACATTTACAGGGCGGCACAGATCGCCAATGCT 180

Qy       59 TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAla 78
            :::|||||||||   |||||||||       ::::::|||:::||||||||||||||
Db      181 CATGACTTCATCCAGGAGTTTCCCAAGGGCTACGATACTGTGGTGGGAGAGAAGGGGGTT 240

Qy       79 GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro 98
               :::|||||||||||||||||||||:::||||||||||||||||||::::::||||||
Db      241 CTGCTGTCAGGTGGACAAAAGCAGAGAGTGGCCATAGCACGAGCCCTGCTGAAGAATCCC 300

Qy       99 LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 118
            |||||||||:::||||||||||||||||||||||||||:::|||::::::         |||
Db      301 AAAATACTTTTACTGGATGAAGCCACGAGTGCTCTGGATGCAGAGAATGAGTTTTTGGTC 360

Qy      119 GlnAlaAlaLeuGluLys 124
                |||      ||||||||:::
Db      361 CAGGAGGCGTTAGAAAGG 378

RESULT 11
AI666988
LOCUS       AI666988                589 bp    mRNA    linear   EST 07-JUN-2001
DEFINITION  fc24c04.y1 Zebrafish WashU MPIMG EST Danio rerio cDNA clone
            IMAGE:3722310 5' similar to TR:Q13040 Q13040 ATP-BINDING CASSETTE
            PROTEIN ;, mRNA sequence.
ACCESSION   AI666988
VERSION     AI666988.1  GI:4805344
KEYWORDS    EST.
SOURCE      zebrafish.
  ORGANISM  Danio rerio
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
            ; Cyprinidae; Danio.
REFERENCE   1  (bases 1 to 589)
  AUTHORS   Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
            ,S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
            ,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
            Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E.,
            Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
            and Wilson,R.
  TITLE     WashU Zebrafish EST Project 1998
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Stephen L. Johnson
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: zbrafish@watson.wustl.edu
            cDNA Library Preparation: Matthew Clark. cDNA Library Arrayed by:
            Matthew Clark. DNA Sequencing by: Washington University Genome
            Sequencing Center Clone distribution: Genome Systems, St. Louis,
            Missouri (web address: www.genomesystems.com) (email contact:
            info@genomesystems.com) and Research Genetics, Huntsville, Alabama
            (web address: www.resgen.com) (email contact: info@resgen.com) and
            RessourcenZentrumPrimarDatenbank, Berlin, Germany (web address:
            www.rzpd.de)
            Seq primer: T3 ET from Amersham
            High quality sequence stop: 481.
FEATURES             Location/Qualifiers
     source          1. .589
                     /organism="Danio rerio"
                     /db_xref="taxon:7955"
                     /clone="IMAGE:3722310"
                     /clone_lib="Zebrafish WashU MPIMG EST"
                     /sex="mixed"
                     /tissue_type="26 somite embryos, adult livers, shield
                     stage embryos"
                     /lab_host="XL1-blue MRF"
                     /note="Vector: pSPORT1; Site_1: NotI; Site_2: SalI; 1st
                     strand cDNA was primed with a Not I - oligo(dT)15 primer
                     [5'pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTTT3'];
                     double-stranded cDNA was ligated to Sal I adaptors (BRL),
                     digested with Not I and cloned into the Not I and Sal I
                     sites of the pSPORT1 vector (BRL). Library was constructed
                     by Matthew Clark (Lehrach lab; ICRF, London and Max Planck
                     Institut fuer Molekulare Genetik,Berlin). cDNAs for EST
                     analysis were selected following oligonucleotide
                     hybridization fingerprinting of arrayed clones from
                     zebrafish late somitogenesis (26 ss), adult liver or
                     embryonic shield stage (5.6 h) libraries.  Fingerprint
                     data were used to computationally cluster cDNAs, and a
                     single cDNA from each cluster was chosen for sequencing.
                     In some cases multiple members of the same cluster were
                     sequenced to assess clustering parameters or single clones
                     were sequenced additional times to assess quality
                     control."
BASE COUNT      151 a    137 c    167 g    133 t      1 others
ORIGIN
```

```
Alignment Scores:
Pred. No.:              8.63e-41      Length:        589
Score:                  378.50        Matches:       79
Percent Similarity:     77.78%        Conservative:  19
Best Local Similarity:  62.70%        Mismatches:    25
Query Match:            57.44%        Indels:        3
DB:                     9             Gaps:          1

US-09-873-409-3 (1-131) x AI666988 (1-589)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
           |||||   :::||||||||  ||||||               |||  ||||||||
Db      1 GTTGATGGTCATGATATTCGAGACTTGAATCCTTATTGGTTAAGGAGTCACATTGGTACT 60

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
          ||||||||||||||||||||||||        :::|||:::  ||||||  ||||||
Db     61 GTAAGCCAGGAACCGGTCCTTTTCTCCTGTTCAATTGCTGAGAACATTGCTTATGGAGCG 120

Qy     42 AspAsp---------ValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla 58
             |||            ||||||  |||::::::  |||||||||:::   |||||||||
Db    121 CCGGATCCGAGTCAGGTGACAGTGGAGGACATTTACAGGGCGGCACAGATCGCCAATGCT 180

Qy     59 TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAla 78
          :::|||||||||   |||||||||       ::::::|||:::|||||||||||||||
Db    181 CATGACTTCATCCAGGAGTTTCCCAAGGGCTACGATACTGTGGTGGGAGAGAAGGGGGTT 240

Qy     79 GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro 98
             :::||||||||||||||||||||:::||||||||||||||||||||::::::||||||
Db    241 CTGCTGTCAGGTGGACAAAAGCAGAGAGTGGCCATAGCACGAGCCCTGCTGAAGAATCCC 300

Qy     99 LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 118
          |||||||||:::||||||||||||||||||||||||||:::|||::::::        |||
Db    301 AAAATACTTTTACTGGATGAAGCCACGAGTGCTCTGGATGCAGAGAATGAGTTTTTGGTC 360

Qy    119 GlnAlaAlaLeuGluLys 124
          |||   |||||||||:::
Db    361 CAGGAGGCGTTAGAAAGG 378

RESULT 12
BJ339421
LOCUS       BJ339421                 542 bp    mRNA    linear   EST 07-MAR-2002
DEFINITION  BJ339421 Dictyostelium discoideum cDNA library, AF Dictyostelium
            discoideum cDNA clone dda65d11 5', mRNA sequence.
ACCESSION   BJ339421
VERSION     BJ339421.1  GI:19247783
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 542)
  AUTHORS   Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
  TITLE     Full length cDNA of Dictyostelium discoideum at the aggregation
            stage
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .542
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="dda65d11"
                     /clone_lib="Dictyostelium discoideum cDNA library, AF"
                     /sex="mat A"
                     /dev_stage="Aggregation stage"
BASE COUNT      186 a     77 c    113 g    166 t
ORIGIN
```

```
Alignment Scores:
Pred. No.:              1.23e-40      Length:        542
Score:                  377.00        Matches:       72
Percent Similarity:     80.70%        Conservative:  20
Best Local Similarity:  63.16%        Mismatches:    22
Query Match:            57.21%        Indels:        0
DB:                     13            Gaps:          0

US-09-873-409-3 (1-131) x BJ339421 (1-542)

Qy      1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
          ::::::|||      |||||||||         |||||||||        |||       ||||||
Db    192 TTATTGGATGGAGAAGATATTAGAAAATTCAATGTCAGAGGGTTACGTCAAAAAATTGGT 251

Qy     21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
          :::|||:::||||||||||||||||||   ||||||||||||   ||||||:::||||||
Db    252 TTAGTCAATCAAGAGCCTGTATTATTCGCCACAACCATTAGTGAAAATATTCGTTATGGT 311

Qy     41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
          ::::::      |||   :::|||:::|||   ||||||:::   |||||||||:::
Db    312 AAAGAAGGTGCCACTCAAGATGAAATTGAAGAAGCAGCTAAATTAGCAAATGCCCATAGT 371

Qy     61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
          ||||||   :::   |||      :::||||||||||||||||||||||||   ||||||
Db    372 TTTATATCACAATTACCACAGGGTTATAATACTTTGGTAGGTGAGAAAGGTGTCCAAATG 431

Qy     81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
          ||||||||||||:::||||||||||||||||||||||||:::::::::||||||    |||
Db    432 TCTGGTGGTCAAAGACAACGTATAGCAATTGCTCGTGCCGTTATTAAAAATCCAAATATA 491

Qy    101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 114
          |||:::|||||||||:::|||||||||||||||:::||||||
Db    492 TTACTCTTGGATGAATCAACATCTGCATTGGATGCAGAGTCT 533

RESULT 13
BJ333166
LOCUS       BJ333166                 547 bp    mRNA    linear   EST 05-MAR-2002
DEFINITION  BJ333166 Dictyostelium discoideum cDNA library, AF Dictyostelium
            discoideum cDNA clone dda41m18 5', mRNA sequence.
ACCESSION   BJ333166
VERSION     BJ333166.1  GI:19163296
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 547)
  AUTHORS   Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
  TITLE     Full length cDNA of Dictyostelium discoideum at the aggregation
            stage
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .547
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="dda41m18"
                     /clone_lib="Dictyostelium discoideum cDNA library, AF"
                     /sex="mat A"
                     /dev_stage="Aggregation stage"
BASE COUNT      186 a     78 c    114 g    169 t
ORIGIN

Alignment Scores:
Pred. No.:              1.24e-40      Length:        547
Score:                  377.00        Matches:       72
Percent Similarity:     80.70%        Conservative:  20
```

```
Best Local Similarity:  63.16%        Mismatches:    22
Query Match:            57.21%        Indels:         0
DB:                     13            Gaps:           0

US-09-873-409-3 (1-131) x BJ333166 (1-547)

Qy        1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
              ::::::|||      ||||||||      |||||||||       |||     ||||||
Db      195 TTATTGGATGGAGAAGATATTAGAAAATTCAATGTCAGAGGGTTACGTCAAAAAAATTGGT 254

Qy       21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
            :::|||:::|||||||||||||||||||    ||||||||||||    |||||||:::||||||
Db      255 TTAGTCAATCAAGAGCCTGTATTATTCGCCACAACCATTAGTGAAAATATTCGTTATGGT 314

Qy       41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
            ::::::      |||   :::|||:::|||    ||||||:::    |||||||||:::
Db      315 AAAGAAGGTGCCACTCAAGATGAAATTGAAGAAGCAGCTAAATTAGCAAATGCCCATAGT 374

Qy       61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
            ||||||   :::   |||         :::||||||||||||||||||||||    ||||||
Db      375 TTTATATCACAATTACCACAGGGTTATAATACTTTGGTAGGTGAGAAAGGTGTCCAAATG 434

Qy       81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
            ||||||||||||:::||||||||||||||||||||||||:::::::::||||||    |||
Db      435 TCTGGTGGTCAAAGACAACGTATAGCAATTGCTCGTGCCGTTATTAAAAATCCAAATATA 494

Qy      101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 114
            |||:::|||||||||:::||||||||||||||||:::||||||
Db      495 TTACTCTTGGATGAATCAACATCTGCATTGGATGCAGAGTCT 536

RESULT 14
BJ356968/c
LOCUS       BJ356968                 589 bp    mRNA    linear   EST 07-MAR-2002
DEFINITION  BJ356968 Dictyostelium discoideum cDNA library, AF Dictyostelium
            discoideum cDNA clone dda62e02 3', mRNA sequence.
ACCESSION   BJ356968
VERSION     BJ356968.1  GI:19256563
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 589)
  AUTHORS   Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
  TITLE     Full length cDNA of Dictyostelium discoideum at the aggregation
            stage
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1..589
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="dda62e02"
                     /clone_lib="Dictyostelium discoideum cDNA library, AF"
                     /sex="mat A"
                     /dev_stage="Aggregation stage"
BASE COUNT      182 a    103 c     99 g    205 t
ORIGIN

Alignment Scores:
Pred. No.:              3.5e-40       Length:       589
Score:                  374.00        Matches:       71
Percent Similarity:     79.51%        Conservative:  26
Best Local Similarity:  58.20%        Mismatches:    25
Query Match:            56.75%        Indels:         0
DB:                     13            Gaps:           0

US-09-873-409-3 (1-131) x BJ356968 (1-589)

Qy        2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
            :::|||   ::::::|||:::   |||||||||||||||    |||            |||:::
Db      541 ATCGATGGTCATAATATAAAGAATCTAAATGTTCGTCATCTTAGACATCTCTTTGGTTTA 482

Qy       22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
               |||   |||||||||    ||||||      ||||||::::::||||||:::||||||:::
Db      481 GTTGGTCAAGAACCAACTCTATTCAGTGGTACAATTGCCGACAATATTCGTTATGGTAAG 422

Qy       42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
               |||   |||   ||||||:::|||   |||::::::   :::|||::::::   |||
Db      421 CATGATGCAACTCAAGAGGAGATTGAAGAGGCTTCAAAATTATCAAATTCTCATTCATTC 362

Qy       62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
            |||::::::   ||||||    :::||||||    :::|||||||||      |||:::|||
Db      361 ATCATTGATTTACCAAATGGTTATAACACAGAGTTGGGTGAGAAATACACTCAATTATCA 302

Qy       82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
            ||||||||||||||||||||||||||||||||||||::::::||||||||||||||||||
Db      301 GGTGGTCAAAAACAACGTATAGCAATTGCTCGTGCAATCATTCGTAATCCAAAGATTTTA 242

Qy      102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
            :::||||||||||:::|||||||||||||||||::::::|||            ||||||   |||
Db      241 TTATTGGATGAGTCAACCTCTGCATTGGATGCTGATTCAACAAAATTAGTTCAAGAAGCA 182

Qy      122 LeuGlu 123
            ||||||
Db      181 CTAGAA 176

RESULT 15
BJ356648/c
LOCUS       BJ356648                 608 bp    mRNA    linear   EST 07-MAR-2002
DEFINITION  BJ356648 Dictyostelium discoideum cDNA library, AF Dictyostelium
            discoideum cDNA clone dda61b02 3', mRNA sequence.
ACCESSION   BJ356648
VERSION     BJ356648.1  GI:19256243
KEYWORDS    EST.
SOURCE      Dictyostelium discoideum.
  ORGANISM  Dictyostelium discoideum
            Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
REFERENCE   1  (bases 1 to 608)
  AUTHORS   Urushihara,H., Tanaka,Y., Kohara,Y. and Shin-i,T.
  TITLE     Full length cDNA of Dictyostelium discoideum at the aggregation
            stage
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1..608
                     /organism="Dictyostelium discoideum"
                     /strain="AX4"
                     /db_xref="taxon:44689"
                     /clone="dda61b02"
                     /clone_lib="Dictyostelium discoideum cDNA library, AF"
                     /sex="mat A"
                     /dev_stage="Aggregation stage"
BASE COUNT      187 a    110 c    101 g    210 t
ORIGIN

Alignment Scores:
Pred. No.:              3.65e-40      Length:       608
Score:                  374.00        Matches:       71
Percent Similarity:     79.51%        Conservative:  26
Best Local Similarity:  58.20%        Mismatches:    25
Query Match:            56.75%        Indels:         0
DB:                     13            Gaps:           0
```

```
US-09-873-409-3 (1-131) x BJ356648 (1-608)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
           :::|||   ::::::|||:::   ||||||||||||||    |||          |||:::
Db     539 ATCGATGGTCATAATATAAAGAATCTAAATGTTCGTCATCTTAGACATCTCTTTGGTTTA 480

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           |||   |||||||||   ||||||         ||||||::::::||||||:::||||||:::
Db     479 GTTGGTCAAGAACCAACTCTATTCAGTGGTACAATTGCCGACAATATTCGTTATGGTAAG 420

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
              |||   |||   ||||||:::|||   |||::::::   :::|||::::::   |||
Db     419 CATGATGCAACTCAAGAGGAGATTGAAGAGGCTTCAAAATTATCAAATTCTCATTCATTC 360

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           |||::::::   ||||||   :::||||||   :::|||||||||      |||:::|||
Db     359 ATCATTGATTTACCAAATGGTTATAACACAGAGTTGGGTGAGAAATACACTCAATTATCA 300

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           ||||||||||||||||||||||||||||||||||||||||||::::::||||||||||||||||||
Db     299 GGTGGTCAAAAACAACGTATAGCAATTGCTCGTGCAATCATTCGTAATCCAAAGATTTTA 240

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
           :::|||||||||:::||||||||||||||::::::|||          ||||||   |||
Db     239 TTATTGGATGAGTCAACCTCTGCATTGGATGCTGATTCAACAAAATTAGTTCAAGAAGCA 180

Qy     122 LeuGlu 123
           ||||||
Db     179 CTAGAA 174

Search completed: March 31, 2003, 13:50:03
Job time : 907.167 secs
```

GenCore version 5.1.4_p5_4578


```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 30, 2003, 01:20:01 ; Search time 128.031 Seconds
                                           (without alignments)
                                           2304.231 Million cell updates/sec

Title:          US-09-873-409-3
Perfect score:  659
Sequence:       1 MVDENDIRALNVRHYRDHIG..........SESKSAVQAALEKDTPRYSF 131

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters:      4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115416_19216/app_query.fasta_1.7544
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_4247_@runat_27032003_115416_19216 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      N_Geneseq_101002:*
                1:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
                2:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
                3:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
                4:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
                5:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
                6:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
                7:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
                8:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
                9:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
                10:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
                11:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
                12:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
                13:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
                14:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
                15:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
                16:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
                17:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
                18:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
                19:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
                20:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
                21:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
                22:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
                23:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
                24:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	619	93.9	3699	24	ABK83223	Human transporter
2	472	71.6	4233	21	AAZ90198	Rat mdr1b2 (multis
3	472	71.6	4233	22	AAF27498	Rat mdr1b2 multidr
4	469	71.2	4279	22	AAD03488	Dog P-glycoprotein
5	469	71.2	4279	22	AAD03504	Dog P-glycoprotein
6	469	71.2	4279	22	AAD03505	Dog P-glycoprotein
7	469	71.2	4279	22	AAD03506	Dog P-glycoprotein
8	468	71.0	4189	21	AAZ49334	Murine multidrug r
9	468	71.0	4189	24	ABA94367	Mouse BCRP DNA rel
10	468	71.0	4313	14	AAQ38950	Mouse multidrug re
11	467	70.9	4254	24	ABK63517	Rat sequence diffe
12	467	70.9	4369	21	AAZ52047	Rat multidrug resi
13	467	70.9	4425	21	AAZ52048	Rat multidrug resi
14	465	70.6	4186	22	AAF86127	Cynomologous monke
15	465	70.6	4195	22	AAF86128	Cynomologous monke
16	465	70.6	4317	22	AAD03489	Dog P-glycoprotein
17	462	70.1	4788	21	AAZ49335	Murine multidrug r
18	462	70.1	4788	24	ABA94368	Mouse BCRP DNA rel
19	461	70.0	3840	24	ABL91687	Human polynucleoti
20	461	70.0	3860	21	AAZ49332	Human wild-type mu
21	461	70.0	3860	21	AAZ49333	Human G185V mutant
22	461	70.0	3860	24	ABA94365	Human BCRP DNA rel
23	461	70.0	3860	24	ABA94366	Human BCRP DNA rel
24	461	70.0	3988	21	AAZ88973	Human MDR-1 DNA.
25	461	70.0	4264	19	AAV66533	Mutated human P-gl
26	461	70.0	4264	19	AAV66534	Mutated human P-gl
27	461	70.0	4349	22	AAH57442	Human intestine ce
28	461	70.0	4378	11	AAQ04522	Multidrug Resistan
29	461	70.0	4646	15	AAQ72872	Human multidrug re
30	461	70.0	4646	21	AAZ94738	Human ATP binding
31	461	70.0	4646	24	AAD38994	Human mdr1 gene.
32	461	70.0	4646	24	ABL68592	Kidney cancer rela
33	461	70.0	4646	24	ABL68880	Kidney cancer rela
34	461	70.0	4669	8	AAN70752	Sequence of human
35	461	70.0	4669	14	AAQ52726	Sequence of human
36	461	70.0	4669	19	AAV32645	Human P glycoprote
37	461	70.0	4669	24	ABK52041	cDNA encoding huma
38	461	70.0	6505	17	AAT13394	Hybrid vector pSF-
39	461	70.0	8630	21	AAZ24041	Retroviral M4 mdr-
40	461	70.0	8630	21	AAZ24042	Retroviral vector
41	448	68.0	3912	24	ABK63653	Rat sequence diffe
42	440	66.8	3924	21	AAZ94742	Human ATP binding
43	440	66.8	3924	21	AAZ88974	Human MDR-3 DNA.
44	440	66.8	3924	24	ABN95801	Gene #2299 used to
45	427	64.8	4776	21	AAZ94744	Human ATP binding

ALIGNMENTS

```
RESULT 1
ABK83223
ID   ABK83223 standard; cDNA; 3699 BP.
XX
AC   ABK83223;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human transporter and ion channel, TRICH14, Incyte ID 7472030CB1, cDNA.
XX
KW   Human; ss; gene; transporter and ion channel; TRICH; transport disorder;
KW   neurological disorder; muscle disorder; immunological disorder; cancer;
KW   scleroderma; systemic lupus erythematosus; allergy; leukaemia;
KW   cell proliferative disorder; cervical cancer; breast cancer;
KW   neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;
KW   myotonic dystrophy; catatonia; endocrine disorder; diabetes;
KW   Grave's disease; gastrointestinal disorder; Crohn's disease;
KW   renal disorder; Good pasture's syndrome; viral infection; cirrhosis;
```

```
KW   bacterial infection; fungal infection; parasitic infection;
KW   protozoal infection; helminthic infection; cardiovascular disorder;
KW   atherosclerosis; hepatic disease.
XX
OS   Homo sapiens.
XX
PN   WO200240541-A2.
XX
PD   23-MAY-2002.
XX
PF   25-OCT-2001; 2001WO-US46055.
XX
PR   27-OCT-2000; 2000US-243989P.
PR   03-NOV-2000; 2000US-245904P.
PR   09-NOV-2000; 2000US-247673P.
PR   17-NOV-2000; 2000US-249661P.
PR   20-NOV-2000; 2000US-252232P.
PR   01-DEC-2000; 2000US-250790P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Tang YT,  Yue H,  Nguyen DB,  Hafalia AJA,  Elliott VS,  Lu Y;
PI   Walia NK,  Yao MG,  Baughn MR,  Gandhi AR,  Ding L,  Sanjanwala M;
PI   Ramkumar J,  Arvizu C,  Gietzen KJ,  Lal PG,  Azimzai Y,  Khan FA;
PI   Thangavelu K,  Thornton M,  Lu DAM,  Tribouley CM,  Warren BA;
PI   Ison CH,  Das D,  Raumann BE,  Policky JL,  Kearney L;
XX
DR   WPI; 2002-463570/49.
DR   P-PSDB; ABG61544.
XX
PT   New transporters and ion channels (TRICH) polypeptides, useful for
PT   diagnosing, preventing, and treating disorders associated with an
PT   abnormal expression or activity of TRICH, e.g. immunological, muscular
PT   or renal disorders  -
XX
PS   Claim 5; Page 172; 178pp; English.
XX
CC   The invention relates to human transporters and ion channels (TRICH)
CC   polypeptides, a naturally occurring amino acid sequence 90 % identical to
CC   TRICH, a biologically active fragment of TRICH or an immunogenic fragment
CC   of TRICH. Also included are an isolated polynucleotide encoding TRICH,
CC   a recombinant polynucleotide comprising a promoter sequence operably
CC   linked to the TRICH polynucleotide, a cell transformed with the
CC   recombinant polynucleotide, a transgenic organism comprising the
CC   recombinant polynucleotide, an isolated antibody that binds specifically
CC   to TRICH, and screening for compounds which bind to TRICH, modulate
CC   TRICH, modulate TRICH expression or are ant/agonists of TRICH.
CC   The polypeptides are useful for diagnosing, treating, and
CC   preventing transport, neurological, muscle, immunological disorders
CC   (e.g. scleroderma, systemic lupus erythematosus, allergies), cell
CC   proliferative disorders such as cancers (e.g. leukaemia, cervical or
CC   breast cancers), neurodegenerative disorders (e.g. Parkinson's disease,
CC   Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy,
CC   catatonia), endocrine disorders (e.g. diabetes, Grave's disease),
CC   gastrointestinal disorders (e.g. Crohn's disease), renal disorders
CC   (e.g. Good pasture's syndrome), viral, bacterial, fungal, parasitic,
CC   protozoal and helminthic infections, cardiovascular disorders (e.g.
CC   atherosclerosis), or hepatic diseases (e.g. cirrhosis) and many
CC   other diseases and disorders detailed in the specification. They can also
CC   be used in assessing the effects of exogenous compounds on the
CC   expression of nucleic acid and amino acid sequences of transporters and
CC   ion channels. TRICH or its fragments may also be used in screening for
CC   compounds that specifically bind to and modulate the activity of TRICH.
CC   The polynucleotides can be used to create knock-in humanised animals or
CC   transgenic animals to model human disease. The present sequence
CC   encodes a TRICH protein.
XX
SQ   Sequence 3699 BP; 1116 A; 707 C; 860 G; 1016 T; 0 other;

Alignment Scores:
Pred. No.:              3.68e-72      Length:          3699
Score:                  619.00        Matches:         124
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            93.93%        Indels:          0
DB:                     24            Gaps:            0

US-09-873-409-3 (1-131) x ABK83223 (1-3699)

Qy       1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1339 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 1398

Qy      21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1399 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 1458

Qy      41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1459 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 1518

Qy      61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1519 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 1578

Qy      81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1579 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 1638

Qy     101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1639 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 1698

Qy     121 AlaLeuGluLys 124
           ||||||||||||
Db    1699 GCACTGGAGAAG 1710

RESULT 2
AAZ90198
ID   AAZ90198 standard; DNA; 4233 BP.
XX
AC   AAZ90198;
XX
DT   19-MAY-2000  (first entry)
XX
DE   Rat mdr1b2 (multispecific drug transporter) nucleotide sequence.
XX
KW   Rat; mdr1b2; multispecific drug transporter; cancer; inflammation;
KW   cardiovascular disease; central nervous system disorder;
KW   autoimmune disease; kidney disease; drug formulation; ss.
XX
OS   Rattus sp.
XX
PN   US6025160-A.
XX
PD   15-FEB-2000.
XX
PF   22-JUL-1998;   98US-0120513.
XX
PR   22-JUL-1998;   98US-0120513.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI   Brun KA,  Ellens H,  Yue L,  Chenery RJ,  Feild JA;
XX
DR   WPI; 2000-181810/16.
DR   P-PSDB; AAY78879.
XX
PT   Isolated polynucleotide encoding a rat mdr1b2 drug transporter
PT   polypeptide, useful for treatment of e.g. cancer, autoimmune disease,
PT   central nervous system disorders  -
XX
PS   Claim 7; Column 13-17; 17pp; English.
XX
CC   This sequence represents a polynucleotide encoding a 1275 amino acid
CC   rat mdr1b2 multispecific drug transporter polypeptide. The mdr1b2
```

```
CC   nucleotide sequence has cytostatic, antiinflammatory, cardiant,
CC   neuroprotective, immunosuppressive and nephrotropic activity.
CC   Understanding the functioning of the mdr1b2 polynucleotide and protein in
CC   transgenic animal models is useful for treating and preventing diseases
CC   such as cancer, inflammation, cardiovascular disease, central nervous
CC   system disorders, autoimmune disease, and kidney disease. The use of the
CC   protein in cell based, membrane based, or binding assays may enhance drug
CC   formulation, selection of formulation excipients and compound design.
XX
SQ   Sequence 4233 BP; 1210 A; 920 C; 1048 G; 1055 T; 0 other;

Alignment Scores:
Pred. No.:              1.96e-52      Length:         4233
Score:                  472.00        Matches:        91
Percent Similarity:     91.06%        Conservative:   21
Best Local Similarity:  73.98%        Mismatches:     11
Query Match:            71.62%        Indels:         0
DB:                     21            Gaps:           0

US-09-873-409-3 (1-131) x AAZ90198 (1-4233)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
            :::|||      ||||||||   :::|||||||||:::    |||:::   |||||||||
Db   1376 ATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTG 1435

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
          ||||||||||||||||||||||||    |||||||||:::    ||||||:::||||||||
Db   1436 GTGAGTCAGGAACCCGTGCTGTTTGCCACCACGATTGCCGAAAACATTCGCTATGGCCGA 1495

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
          ::::::||||||   :::|||:::|||:::|||   :::|||||||||||||||||||||
Db   1496 GAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTC 1555

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
          ||||||:::   |||:::|||||||||||||||||||||||||:::|||||||||:::|||
Db   1556 ATCATGAAACTGCCCCACAAATTTAACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGT 1615

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1616 GGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAACCCCAAGATCCTT 1675

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
          :::|||||||||||||||||||||||||||:::||||||::::::   ||||||||||||
Db   1676 TTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTGGTTCAGGCCGCT 1735

Qy    122 LeuGluLys 124
          |||:::|||
Db   1736 CTGGATAAG 1744

RESULT 3
AAF27498
ID   AAF27498 standard; cDNA; 4233 BP.
XX
AC   AAF27498;
XX
DT   25-APR-2001  (first entry)
XX
DE   Rat mdr1b2 multidrug resistance cDNA, SEQ ID NO:1.
XX
KW   Rat; multidrug resistance; mdr1b2; multi-drug transporter family;
KW   drug screening; pharmacokinetic analysis; oral absorption;
KW   formulation design; bioavailability; transgenic animal; knockout animal;
KW   inflammation; cardiovascular disease; central nervous system disorder;
KW   cancer; autoimmune disorder; kidney disease; ss.
XX
OS   Rattus sp.
XX
PN   US6169166-B1.
XX
PD   02-JAN-2001.
XX
PF   29-NOV-1999;   99US-0450105.
XX
PR   22-JUL-1998;   98US-0120513.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI   Chenery RJ,  Ellens H,  Brun KA,  Yue L,  Feild JA;
XX
DR   WPI; 2001-158280/16.
DR   P-PSDB; AAB60409.
XX
PT   New rat multidrug resistance protein (mdr1b2), useful for screening or
PT   identifying compounds that are (ant)agonists/inhibitors of the
PT   ratmdr1b2, as well as compounds with optimal development
PT   characteristics  -
XX
PS   Disclosure; Column 13-14; 13pp; English.
XX
CC   The invention relates to a novel rat multidrug resistance (mdr)
CC   protein, mdr1b2 (AAB60409). Multi-specific drug transporter family
CC   proteins are present in cell which have a barrier function, such as
CC   intestinal epithelial cells, brain microvessel endothelial cells,
CC   kidney epithelial cells, and liver hepatocytes, and are also expressed
CC   by certain cancer cells. The rat mdr1b2 protein is useful for
CC   screening or identifying compounds that are agonists or antagonists
CC   of mdr1b2 activity. It may also be used to establish assays to
CC   predict oral absorption and pharmacokinetics of drugs in humans, and
CC   thus enhance the design of formulations through the identification of
CC   compounds with optimal development characteristics (i.e., high oral
CC   bioavailability, UID (once a day) dosing, reduced drug interactions,
CC   reduced variability, and reduced food effects), specifically to
CC   avoid interactions with human mdr-1. Transgenic and knockout animals
CC   created using DNA encoding the rat mdr1b2 may be used to gain an
CC   insight into treating and preventing human diseases such as cancer,
CC   inflammation, cardiovascular disease, central nervous system disorders,
CC   autoimmune disorders and kidney disease. The present sequence represents
CC   cDNA encoding rat mdr1b2.
XX
SQ   Sequence 4233 BP; 1210 A; 920 C; 1048 G; 1055 T; 0 other;

Alignment Scores:
Pred. No.:              1.96e-52      Length:         4233
Score:                  472.00        Matches:        91
Percent Similarity:     91.06%        Conservative:   21
Best Local Similarity:  73.98%        Mismatches:     11
Query Match:            71.62%        Indels:         0
DB:                     22            Gaps:           0

US-09-873-409-3 (1-131) x AAF27498 (1-4233)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
            :::|||      ||||||||   :::|||||||||:::    |||:::   |||||||||
Db   1376 ATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTG 1435

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
          ||||||||||||||||||||||||    |||||||||:::    ||||||:::||||||||
Db   1436 GTGAGTCAGGAACCCGTGCTGTTTGCCACCACGATTGCCGAAAACATTCGCTATGGCCGA 1495

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
          ::::::||||||   :::|||:::|||:::|||   :::|||||||||||||||||||||
Db   1496 GAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTC 1555

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
          ||||||:::   |||:::|||||||||||||||||||||||||:::|||||||||:::|||
Db   1556 ATCATGAAACTGCCCCACAAATTTAACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGT 1615

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1616 GGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAACCCCAAGATCCTT 1675

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
          :::|||||||||||||||||||||||||||:::||||||::::::   ||||||||||||
Db   1676 TTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTGGTTCAGGCCGCT 1735
```

```
Qy      122 LeuGluLys 124
            |||:::|||
Db     1736 CTGGATAAG 1744

RESULT 4
AAD03488
ID   AAD03488 standard; cDNA; 4279 BP.
XX
AC   AAD03488;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) cDNA #1.
XX
KW   Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag=  a
FT                   /product= "Dog P-glycoprotein (PGP) #1"
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00303.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 3; Page 58-63; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) cDNA. This
CC   sequence is also referred as Genotype C cDNA. The
CC   PGP enzyme functions as an efflux pump exporting small molecules
CC   across the cell membrane. This enzyme is a member of the ABC
CC   transporter family.
XX
SQ   Sequence 4279 BP; 1294 A; 834 C; 1008 G; 1143 T; 0 other;

Alignment Scores:
Pred. No.:              5e-52          Length:         4279
Score:                  469.00         Matches:        90
Percent Similarity:     89.43%         Conservative:   20
Best Local Similarity:  73.17%         Mismatches:     13
Query Match:            71.17%         Indels:         0
DB:                     22             Gaps:           0

US-09-873-409-3 (1-131) x AAD03488 (1-4279)

Qy        2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
            :::|||      |||||||||   :::||||||||||||   |||:::      ||||||
Db     1379 ATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTG 1438

Qy       22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
            ||||||||||||||||||||||||   |||||||||:::   ||||||:::|||||||||
Db     1439 GTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGC 1498

Qy       42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
            ::::::||||||   :::|||:::|||:::|||   :::|||||||||||||||||||||
Db     1499 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTT 1558

Qy       62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
            ||||||:::   ||||||||||||:::|||||||||||||||:::|||||||||:::|||
Db     1559 ATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGT 1618

Qy       82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1619 GGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTT 1678

Qy      102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
            :::|||||||||||||||||||||||||||:::||||||::::::   ||||||   |||
Db     1679 CTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCC 1738

Qy      122 LeuGluLys 124
            |||:::|||
Db     1739 CTGGATAAG 1747

RESULT 5
AAD03504
ID   AAD03504 standard; cDNA; 4279 BP.
XX
AC   AAD03504;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) allelic variant (Genotype A) cDNA.
XX
KW   Dog; P-glycoprotein allelic variant; multidrug transporter; MDR1;
KW   drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag=  a
FT                   /product= "Dog P-glycoprotein (PGP) allelic variant
FT                   (Genotype A) protein"
FT   allele          replace (607, C)
FT                   /*tag=  b
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00308.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
```

```
PS   Claim 9; Page 85-90; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) allelic variant
CC   (Genotype A) cDNA. The PGP enzyme functions as an efflux pump
CC   exporting small molecules across the cell membrane. This enzyme
CC   is a member of the ABC transporter family.
XX
SQ   Sequence 4279 BP; 1295 A; 833 C; 1008 G; 1143 T; 0 other;

Alignment Scores:
Pred. No.:              5e-52          Length:        4279
Score:                  469.00         Matches:       90
Percent Similarity:     89.43%         Conservative:  20
Best Local Similarity:  73.17%         Mismatches:    13
Query Match:            71.17%         Indels:        0
DB:                     22             Gaps:          0

US-09-873-409-3 (1-131) x AAD03504 (1-4279)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
          :::||      ||||||||   :::|||||||||||   |||:::       |||||
Db   1379 ATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTG 1438

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
          ||||||||||||||||||||||||   |||||||||:::   ||||||:::|||||||||
Db   1439 GTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGC 1498

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
          ::::::||||||   :::|||:::|||:::|||   :::|||||||||||||||||||||
Db   1499 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTT 1558

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
          ||||||:::   ||||||||||||:::||||||||||||||||:::|||||||||:::|||
Db   1559 ATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGT 1618

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1619 GGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTT 1678

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
          :::||||||||||||||||||||||||||||||:::||||||::::::   ||||||   |||
Db   1679 CTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCC 1738

Qy    122 LeuGluLys 124
          |||:::|||
Db   1739 CTGGATAAG 1747

RESULT 6
AAD03505
ID   AAD03505 standard; cDNA; 4279 BP.
XX
AC   AAD03505;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.
XX
KW   Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KW   MDR1; drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag=  a
FT                   /product= "Dog P-glycoprotein (PGP) allelic variant
FT                   (Genotype B) protein"
FT   allele          replace (91, T)
FT                   /*tag=  b
FT   allele          replace (607, C)
FT                   /*tag=  c
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00309.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 9; Page 93-99; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) allelic variant
CC   (Genotype B) cDNA. The PGP enzyme functions as an efflux pump
CC   exporting small molecules across the cell membrane. This enzyme
CC   is a member of the ABC transporter family.
XX
SQ   Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;

Alignment Scores:
Pred. No.:              5e-52          Length:        4279
Score:                  469.00         Matches:       90
Percent Similarity:     89.43%         Conservative:  20
Best Local Similarity:  73.17%         Mismatches:    13
Query Match:            71.17%         Indels:        0
DB:                     22             Gaps:          0

US-09-873-409-3 (1-131) x AAD03505 (1-4279)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
          :::||      ||||||||   :::|||||||||||   |||:::       |||||
Db   1379 ATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTG 1438

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
          ||||||||||||||||||||||||   |||||||||:::   ||||||:::|||||||||
Db   1439 GTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGC 1498

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
          ::::::||||||   :::|||:::|||:::|||   :::|||||||||||||||||||||
```

```
Db   1499 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTT 1558

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
          ||||||:::   |||||||||||:::|||||||||||||||:::|||||||||:::|||
Db   1559 ATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGT 1618

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1619 GGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTT 1678

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
          :::|||||||||||||||||||||||||:::||||||:::::   ||||||    |||
Db   1679 CTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCC 1738

Qy    122 LeuGluLys 124
          |||:::|||
Db   1739 CTGGATAAG 1747
```

```
RESULT 7
AAD03506
ID   AAD03506 standard; cDNA; 4279 BP.
XX
AC   AAD03506;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) allelic variant (Genotype D) cDNA.
XX
KW   Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KW   MDR1; drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag= a
FT                   /product= "Dog P-glycoprotein (PGP) allelic variant
FT                   (Genotype D) protein"
FT   allele          replace (91, T)
FT                   /*tag= b
FT   allele          replace (607, C)
FT                   /*tag= c
FT   allele          replace (1001, T)
FT                   /*tag= c
FT   allele          replace (3458, A)
FT                   /*tag= c
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00310.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 9; Page 102-107; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) allelic variant
CC   (Genotype D) cDNA. The PGP enzyme functions as an efflux pump
CC   exporting small molecules across the cell membrane. This enzyme
CC   is a member of the ABC transporter family.
XX
SQ   Sequence 4279 BP; 1296 A; 833 C; 1009 G; 1141 T; 0 other;

Alignment Scores:
Pred. No.:              5e-52          Length:        4279
Score:                  469.00         Matches:       90
Percent Similarity:     89.43%         Conservative:  20
Best Local Similarity:  73.17%         Mismatches:    13
Query Match:            71.17%         Indels:        0
DB:                     22             Gaps:          0

US-09-873-409-3 (1-131) x AAD03506 (1-4279)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
          :::|||      ||||||||   :::|||||||||||   |||:::         ||||||
Db   1379 ATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTG 1438

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
          ||||||||||||||||||||||||   |||||||||:::   ||||||:::|||||||||
Db   1439 GTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGC 1498

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
          ::::::||||||   :::|||:::|||:::|||   :::||||||||||||||||||||||
Db   1499 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTT 1558

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
          ||||||:::   |||||||||||:::|||||||||||||||:::|||||||||:::|||
Db   1559 ATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGT 1618

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1619 GGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTT 1678

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
          :::|||||||||||||||||||||||||:::||||||:::::   ||||||    |||
Db   1679 CTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCC 1738

Qy    122 LeuGluLys 124
          |||:::|||
Db   1739 CTGGATAAG 1747
```

```
RESULT 8
AAZ49334
ID   AAZ49334 standard; cDNA; 4189 BP.
XX
AC   AAZ49334;
XX
DT   14-MAR-2000  (first entry)
XX
DE   Murine multidrug resistance-1 (MDR-1) cDNA.
XX
KW   Multidrug resistance; MDR-1; P-glycoprotein;
KW   transmembrane efflux pump; haematopoietic stem cell; transduction;
KW   bone marrow transplantation; chemotherapy; radiation therapy; cancer;
KW   gene therapy; gene replacement; genetic defect; thalassaemia;
KW   Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
KW   cytokine; ds.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
```

```
FT   CDS             1..3831
FT                   /*tag=  a
FT                   /product= "Murine MDR-1 protein"
XX
PN   WO9961589-A2.
XX
PD   02-DEC-1999.
XX
PF   27-MAY-1999;   99WO-US11825.
XX
PR   28-MAY-1998;   98US-0086988.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI   Sorrentino B,  Bunting K;
XX
DR   WPI; 2000-072615/06.
DR   P-PSDB; AAY58188.
XX
PT   Ex vivo expansion of hematopoietic stem cells transduced with a
PT   sequence encoding human multidrug resistance-1, used for bone marrow
PT   transplantation -
XX
PS   Disclosure; Page 90-93; 113pp; English.
XX
CC   This sequence represents cDNA encoding murine multidrug
CC   resistance protein MDR-1. MDR-1 is a transmembrane
CC   efflux pump, responsible for the export of drugs from cells,
CC   particularly cancer cells. The invention relates to transducing
CC   haematopoietic stem cells with nucleic acid encoding an MDR protein
CC   and culturing the modified cells. The modified haematopoietic stem
CC   cells are useful in bone marrow transplantation (to reconstitute
CC   haematopoietic systems in patients who have undergone chemotherapy or
CC   radiation therapy) and in ex vivo gene therapy of genetic defects in
CC   cells derived from haematopoietic stem cells, e.g., thalassaemia,
CC   Gauchier's disease, sickle cell anaemia or leukaemia. The modified
CC   cells can also be used to identify factors involved in regulating
CC   proliferation and differentiation in haematopoietic stem cells.
CC   Haematopoietic stem cells that express MDR-1 will be protected against
CC   chemotherapeutic agents, so can be engrafted while the patient is
CC   undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells
CC   provides sufficient cells to permit standard biochemical analysis.
CC   Overexpression of MDR-1 allows cytokine-driven expansion of
CC   haematopoietic stem cells by at least 10-fold compared with a maximum
CC   of 4-fold in known procedures.
XX
SQ   Sequence 4189 BP; 1204 A; 875 C; 1028 G; 1082 T; 0 other;

Alignment Scores:
Pred. No.:              6.6e-52       Length:        4189
Score:                  468.00        Matches:       90
Percent Similarity:     91.06%        Conservative:  22
Best Local Similarity:  73.17%        Mismatches:    11
Query Match:            71.02%        Indels:        0
DB:                     21            Gaps:          0

US-09-873-409-3 (1-131) x AAZ49334 (1-4189)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
          :::|||      ||||||||     :::||||||||:::    |||:::    |||||||||
Db   1354 ATCGACGGACAAGACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCATTGGTGTG 1413

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
          ||||||||||||||||||||||||||||||     ||||||||||:::    |||||:::||||||||||
Db   1414 GTGAGTCAGGAACCTGTGCTGTTTGCCACCACGATCGCCGAGAACATTCGCTATGGCCGA 1473

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
          :::|||||||||    :::|||:::|||:::|||    :::||||||||||||||||||||||||
Db   1474 GAAGATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTC 1533

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
          ||||||:::    |||::::::|||:::|||||||||||||||:::|||||||||:::|||
Db   1534 ATCATGAAACTGCCCCACCAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGT 1593

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1594 GGGGGACAGAAACAGAGAATCGCCATTGCCCGGGCCCTGGTCCGCAATCCCAAGATCCTT 1653

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
          :::|||||||||||||||||||||||||||:::||||||::::::    ||||||||||||
Db   1654 TTGTTGGACGAGGCCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTGGTGCAGGCCGCA 1713

Qy    122 LeuGluLys 124
          |||:::|||
Db   1714 CTGGATAAG 1722

RESULT 9
ABA94367
ID   ABA94367 standard; DNA; 4189 BP.
XX
AC   ABA94367;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Mouse BCRP DNA related seq Id No. 5.
XX
KW   Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;
KW   hepatotropic; neurodegenerative; cytostatic; antianemic; muscular; BCRP;
KW   cardiant; gene therapy; ds.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3831
FT                   /*tag=  a
XX
PN   WO200192877-A2.
XX
PD   06-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US17459.
XX
PR   31-MAY-2000; 2000US-0584586.
PR   29-MAY-2001; 2001US-0866866.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI   Sorrentino B,  Schuetz J;
XX
DR   WPI; 2002-114368/15.
DR   P-PSDB; ABB07268.
XX
PT   Identifying a stem cell, for treating e.g., muscular dystrophy,
PT   myocardial infarction, Parkinson's disease, or neurodegenerative
PT   disorders, comprises detecting the expression of an ATP transport
PT   protein (BCRP) by a cell  -
XX
PS   Disclosure; Page 64-66; 87pp; English.
XX
CC   The invention provides a method of identifying and/or isolating a stem
CC   cell that involves detecting the expression of an ATP transport protein
CC   containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
CC   comprising stem cells. The isolated stem cells may be used in the
CC   treatment of diseases such as muscular dystrophy, degenerative liver
CC   disorder, myocardial infarction, Parkinson's disease, degenerative
CC   disorders of the brain, and for tissue regeneration or replacement.
CC   Haematopoietic cells can be used in bone marrow transplants (e.g., for
CC   treatment of leukemia) and for ex vivo gene therapy for treating blood
CC   diseases such as sickle cell anemia and thalassemia. The stem cells can
CC   also be used as cell targets in gene therapy protocols. The present
CC   sequence represents a sequence related to the BCRP for which no relevant
CC   information has been provided in the specification.
XX
SQ   Sequence 4189 BP; 1204 A; 875 C; 1028 G; 1082 T; 0 other;

Alignment Scores:
```

```
Pred. No.:              6.6e-52      Length:         4189
Score:                  468.00       Matches:        90
Percent Similarity:     91.06%       Conservative:   22
Best Local Similarity:  73.17%       Mismatches:     11
Query Match:            71.02%       Indels:         0
DB:                     24           Gaps:           0

US-09-873-409-3 (1-131) x ABA94367 (1-4189)

Qy        2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
            :::|||      ||||||||   :::|||||||||:::   |||:::   ||||||||
Db     1354 ATCGACGGACAAGACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCATTGGTGTG 1413

Qy       22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
            ||||||||||||||||||||||   |||||||||:::   ||||||:::||||||||
Db     1414 GTGAGTCAGGAACCTGTGCTGTTTGCCACCACGATCGCCGAGAACATTCGCTATGGCCGA 1473

Qy       42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
            :::|||||||||   :::|||:::|||:::|||   :::||||||||||||||||||||||
Db     1474 GAAGATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTC 1533

Qy       62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
            ||||||:::   |||::::::|||:::|||||||||||||||:::|||||||||:::|||
Db     1534 ATCATGAAACTGCCCCACCAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGT 1593

Qy       82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1594 GGGGGACAGAAACAGAGAATCGCCATTGCCCGGGCCCTGGTCCGCAATCCCAAGATCCTT 1653

Qy      102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
            :::|||||||||||||||||||||||||||:::||||||::::::   |||||||||||
Db     1654 TTGTTGGACGAGGCCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTGGTGCAGGCCGCA 1713

Qy      122 LeuGluLys 124
            |||:::|||
Db     1714 CTGGATAAG 1722
```

```
RESULT 10
AAQ38950
ID   AAQ38950 standard; DNA; 4313 BP.
XX
AC   AAQ38950;
XX
DT   28-JUL-1993  (first entry)
XX
DE   Mouse multidrug resistance sequence.
XX
KW   mdr gene; Lambda DR11 clone; ss.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   CDS             110..3940
FT                   /*tag=  a
FT                   /phenotype= multidrug_resistance
XX
PN   US5198344-A.
XX
PD   30-MAR-1993.
XX
PF   15-JUL-1986;   86US-0885951.
XX
PR   15-JUL-1986;   86US-0885951.
PR   06-FEB-1991;   91US-0652311.
XX
PA   (MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
PI   Croop JM,  Gros P,  Housman DE;
XX
DR   WPI; 1993-126077/15.
DR   P-PSDB; AAR35199.
XX
PT   DNA sequence which confers multi-drug resistance on sensitive
PT   mammalian cells - used to preserve bone marrow cells during
PT   chemotherapy to prevent infection
XX
PS   Claim 1; Fig 8; 22pp; English.
XX
CC   A cDNA library was constructed from mRNA which had been isolated
CC   from a drug-sensitive mouse cell-line. Two mouse cDNA molecules
CC   complementary to the mRNA species encoded by 2 related but distinct
CC   mdr genes were isolated and cloned. One of the cDNA clones (lambda
CC   DR11) is a full-length cDNA clone for one member of the mdr gene
CC   family. When incorporated into prokaryotic expression vector pDREX4
CC   (which allows high levels of transcription of the cDNA when
CC   introduced into mammalian cells), the clone was shown to confer the
CC   multidrug resistance phenotype upon transfection into drug-sensitive
CC   mammalian cells.
XX
SQ   Sequence 4313 BP; 1241 A; 904 C; 1058 G; 1110 T; 0 other;

Alignment Scores:
Pred. No.:              6.87e-52     Length:         4313
Score:                  468.00       Matches:        90
Percent Similarity:     91.06%       Conservative:   22
Best Local Similarity:  73.17%       Mismatches:     11
Query Match:            71.02%       Indels:         0
DB:                     14           Gaps:           0

US-09-873-409-3 (1-131) x AAQ38950 (1-4313)

Qy        2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
            :::|||      ||||||||   :::|||||||||:::   |||:::   ||||||||
Db     1463 ATCGACGGACAAGACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCATTGGTGTG 1522

Qy       22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
            ||||||||||||||||||||||   |||||||||:::   ||||||:::||||||||
Db     1523 GTGAGTCAGGAACCTGTGCTGTTTGCCACCACGATCGCCGAGAACATTCGCTATGGCCGA 1582

Qy       42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
            :::|||||||||   :::|||:::|||:::|||   :::||||||||||||||||||||||
Db     1583 GAAGATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTC 1642

Qy       62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
            ||||||:::   |||::::::|||:::|||||||||||||||:::|||||||||:::|||
Db     1643 ATCATGAAACTGCCCCACCAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGT 1702

Qy       82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1703 GGGGGACAGAAACAGAGAATCGCCATTGCCCGGGCCCTGGTCCGCAATCCCAAGATCCTT 1762

Qy      102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
            :::|||||||||||||||||||||||||||:::||||||::::::   |||||||||||
Db     1763 TTGTTGGACGAGGCCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTGGTGCAGGCCGCA 1822

Qy      122 LeuGluLys 124
            |||:::|||
Db     1823 CTGGATAAG 1831
```

```
RESULT 11
ABK63517
ID   ABK63517 standard; cDNA; 4254 BP.
XX
AC   ABK63517;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Rat sequence differentially expressed in response to a hepatotoxin #1424.
XX
KW   Rat; ss; hepatotoxin; expressed sequence tag; EST; drug screening;
KW   differential expression; centrilobular necrosis; steatosis.
XX
OS   Rattus norvegicus.
XX
PN   WO200210453-A2.
XX
```

```
PD   07-FEB-2002.
XX
PF   30-JUL-2001; 2001WO-US23872.
XX
PR   31-JUL-2000; 2000US-222040P.
PR   02-NOV-2000; 2000US-244880P.
PR   11-MAY-2001; 2001US-290029P.
PR   15-MAY-2001; 2001US-290645P.
PR   22-MAY-2001; 2001US-292336P.
PR   06-JUN-2001; 2001US-295798P.
PR   13-JUN-2001; 2001US-297457P.
PR   19-JUN-2001; 2001US-298884P.
PR   09-JUL-2001; 2001US-303459P.
XX
PA   (GENE-) GENE LOGIC INC.
XX
PI   Mendrick D,  Porter MW,  Johnson KR,  Castle AL,  Elashoff MR;
XX
DR   WPI; 2002-241625/29.
XX
PT   Predicting toxic effects of compounds or the progression of these toxic
PT   effects by determining the changes in gene expression in tissues or
PT   cells exposed to the toxin and comparing these to gene expression in
PT   unexposed tissues or cells -
XX
PS   Claim 1; Seq ID No 1424; 239pp; English.
XX
CC   The invention relates to methods for predicting toxic effects of
CC   compounds or the progression of these toxic effects by determining the
CC   global changes in gene expression in tissues or cells exposed to the
CC   toxin and comparing these to gene expression in unexposed tissues or
CC   cells. Also included are methods of predicting at least one toxic
CC   effect of a compound or progression of a toxic effect, preferably the
CC   hepatotoxicity of a compound, comprising detecting the level of
CC   expression in a tissue or cell sample exposed to the compound of two or
CC   more genes listed in the specification, where  differential expression of
CC   the genes is indicative of at least one toxic effect or progression.
CC   The method can also be used to identify an agent which modulates the
CC   toxic response and predict cellular pathways that a compound modulates
CC   in a cell. The methods utilise a set of at least two probes (on a solid
CC   support in kit form), where each of the probes comprises a sequence that
CC   specifically hybridises to a gene listed in the specification, a computer
CC   system comprising a database containing information identifying the
CC   expression level in a tissue or cell sample exposed to a hepatotoxin of a
CC   set of genes comprising at least two genes listed in the specification,
CC   and a user interface to view the information used to present information
CC   identifying the expression level in a tissue or cell of at least one gene
CC   listed in the specification. The method is useful for elucidating global
CC   changes in gene expression and for identifying toxicity markers in
CC   tissues or cell exposed to a known toxin. The genes may be used as
CC   toxicity markers in drug screening and toxicity assays. The genes and
CC   gene expression information may be used as diagnostic markers for the
CC   prediction or identification of the physiological state of tissue or cell
CC   sample that has been exposed to a compound or agent. Hepatotoxicity
CC   is characterised by centrilobular necrosis and steatosis. The present
CC   sequence is an expressed sequence tag (EST) or cDNA derived from a gene
CC   which is differentially expressed in response to a hepatotoxic agent.
XX
SQ   Sequence 4254 BP; 1188 A; 934 C; 1066 G; 1066 T; 0 other;

Alignment Scores:
Pred. No.:              9.17e-52       Length:        4254
Score:                  467.00         Matches:       90
Percent Similarity:     91.06%         Conservative:  22
Best Local Similarity:  73.17%         Mismatches:    11
Query Match:            70.86%         Indels:        0
DB:                     24             Gaps:          0

US-09-873-409-3 (1-131) x ABK63517 (1-4254)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
          :::|||      |||||||||   :::|||||||||:::   |||:::   |||||||||
Db   1456 ATTGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTG 1515

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
          ||||||||||||||||||||||||   |||||||||:::   ||||||:::|||||||||
Db   1516 GTGAGTCAGGAACCCGTGCTGTTTGCCACCACGATTGCCGAAAACATTCGCTATGGCCGA 1575

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
          ::::::||||||   :::|||:::|||:::|||   :::|||||||||||||||||||||
Db   1576 GAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTC 1635

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
          ||||||:::   |||:::||||||:::|||||||||||||||:::|||||||||:::|||
Db   1636 ATCATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGT 1695

Qy     82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1696 GGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAACCCCAAGATCCTT 1755

Qy    102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
          :::|||||||||||||||||||||||||||:::||||||::::::   ||||||||||||
Db   1756 TTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTGGTTCAGGCCGCT 1815

Qy    122 LeuGluLys 124
          |||:::|||
Db   1816 CTGGATAAG 1824

RESULT 12
AAZ52047
ID   AAZ52047 standard; cDNA; 4369 BP.
XX
AC   AAZ52047;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Rat multidrug resistance protein 1a cDNA.
XX
KW   Multidrug resistance protein 1a; mdr1a; multi-specific drug transporter;
KW   drug formulation; formulation excipient; compound design; inflammation;
KW   pharmacokinetic; oral absorption; cancer; cardiovascular disease;
KW   central nervous system disorder; auto-immune disease; kidney disease; ss.
XX
OS   Rattus rattus.
XX
FH   Key             Location/Qualifiers
FT   CDS             352..4170
FT                   /*tag=  a
FT                   /product= "Rat multidrug resistance protein"
XX
PN   WO200015650-A1.
XX
PD   23-MAR-2000.
XX
PF   10-SEP-1999;   99WO-US20770.
XX
PR   17-SEP-1998;   98US-0156800.
PR   09-DEC-1998;   98US-0208809.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI   Brun KA,  Chenery RJ,  Ellens H,  Feild JA,  Yue L;
XX
DR   WPI; 2000-271372/23.
DR   P-PSDB; AAY70596.
XX
PT   Isolated rat mdr1a polynucleotides and polypeptides, useful in assays
PT   to provide information on drug formulation, selection of formulation
PT   excipients and compound design -
XX
PS   Claim 2; Page 23-24; 33pp; English.
XX
CC   The present cDNA sequence encodes rat multidrug resistance
CC   protein 1a (mdr1a). This protein is a member of the multi-specific drug
CC   transporters family. Mdr1a is used in assays to provide information on
CC   drug formulation, selection of formulation excipients and compound
```

```
CC   design. They are used in cell based, membrane based, binding or other
CC   assays to provide information that may enhance drug formulation. This
CC   invention further relates to the generation of in vivo and in vitro
CC   comparison data to predict oral absorption and pharmacokinetics. This
CC   enables the selection of drugs with optimal pharmacokinetics, i.e. good
CC   oral bioavailability, brain penetration, plasma half life, and minimum
CC   drug interaction. Transgenic and knock-out animals created using rat
CC   mdr1a provides an insight into treating and preventing human diseases
CC   including cancer, inflammation, cardiovascular disease, central nervous
CC   system disorders, auto-immune and kidney disease.
XX
SQ   Sequence 4369 BP; 1241 A; 973 C; 1096 G; 1059 T; 0 other;

Alignment Scores:
Pred. No.:              9.51e-52      Length:         4369
Score:                  467.00        Matches:        90
Percent Similarity:     91.06%        Conservative:   22
Best Local Similarity:  73.17%        Mismatches:     11
Query Match:            70.86%        Indels:         0
DB:                     21            Gaps:           0

US-09-873-409-3 (1-131) x AAZ52047 (1-4369)

Qy        2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
              :::|||       ||||||||   :::|||||||||:::   |||:::   |||||||||
Db     1684 ATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTG 1743

Qy       22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
            ||||||||||||||||||||||||    |||||||||:::   ||||||:::|||||||||
Db     1744 GTGAGTCAGGAACCCGTGCTGTTTGCCACCACAATTGCCGAAAACATTCGCTATGGCCGA 1803

Qy       42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
            ::::::||||||   :::|||:::|||:::|||   :::||||||||||||||||||||||
Db     1804 GAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAATGCCTATGATTTC 1863

Qy       62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
            ||||||:::   |||:::||||||:::|||||||||||||||:::|||||||||:::|||
Db     1864 ATCATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGT 1923

Qy       82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1924 GGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAACCCCAAGATCCTT 1983

Qy      102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
            :::||||||||||||||||||||||||||:::||||||::::::   ||||||||||||
Db     1984 TTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTGGTTCAGGCCGCT 2043

Qy      122 LeuGluLys 124
            |||:::|||
Db     2044 CTGGATAAG 2052

RESULT 13
AAZ52048
ID   AAZ52048 standard; cDNA; 4425 BP.
XX
AC   AAZ52048;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Rat multidrug resistance protein 1a cDNA derived from EST sequences.
XX
KW   Multidrug resistance protein 1a; mdr1a; multi-specific drug transporter;
KW   drug formulation; formulation excipient; compound design; inflammation;
KW   pharmacokinetic; oral absorption; cancer; cardiovascular disease;
KW   central nervous system disorder; auto-immune disease; kidney disease;
KW   EST; expressed sequence tag; ss.
XX
OS   Rattus rattus.
XX
FH   Key             Location/Qualifiers
FT   CDS             352..4170
FT                   /*tag=  a
FT                   /product= "Rat multidrug resistance protein"
XX
PN   WO200015650-A1.
XX
PD   23-MAR-2000.
XX
PF   10-SEP-1999;   99WO-US20770.
XX
PR   17-SEP-1998;   98US-0156800.
PR   09-DEC-1998;   98US-0208809.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI   Brun KA,  Chenery RJ,  Ellens H,  Feild JA,  Yue L;
XX
DR   WPI; 2000-271372/23.
DR   P-PSDB; AAY70597.
XX
PT   Isolated rat mdr1a polynucleotides and polypeptides, useful in assays
PT   to provide information on drug formulation, selection of formulation
PT   excipients and compound design -
XX
PS   Claim 11; Page 27-29; 33pp; English.
XX
CC   The present cDNA sequence encodes rat multidrug resistance protein 1a
CC   (mdr1a). This cDNA is derived from EST (expressed sequence tag)
CC   sequences. Mdr1a is used in assays to provide information on
CC   drug formulation, selection of formulation excipients and compound
CC   design. They are used in cell based, membrane based, binding or other
CC   assays to provide information that may enhance drug formulation. This
CC   invention further relates to the generation of in vivo and in vitro
CC   comparision data to predict oral absorption and pharmacokinetics. This
CC   enables the selection of drugs with optimal pharmacokinetics, i.e. good
CC   oral bioavailability, brain penetration, plasma half life, and minimum
CC   drug interaction. Transgenic and knock-out animals created using rat
CC   mdr1a provides an insight into treating and preventing human diseases
CC   including cancer, inflammation, cardiovascular disease, central nervous
CC   system disorders, auto-immune and kidney disease.
XX
SQ   Sequence 4425 BP; 1250 A; 992 C; 1113 G; 1067 T; 2 U; 1 other;

Alignment Scores:
Pred. No.:              9.68e-52      Length:         4425
Score:                  467.00        Matches:        90
Percent Similarity:     91.06%        Conservative:   22
Best Local Similarity:  73.17%        Mismatches:     11
Query Match:            70.86%        Indels:         0
DB:                     21            Gaps:           0

US-09-873-409-3 (1-131) x AAZ52048 (1-4425)

Qy        2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
              :::|||       ||||||||   :::|||||||||:::   |||:::   |||||||||
Db     1684 ATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTG 1743

Qy       22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
            ||||||||||||||||||||||||    |||||||||:::   ||||||:::|||||||||
Db     1744 GTGAGTCAGGAACCCGTGCTGTTTGCCACCACAATTGCCGAAAACATTCGCTATGGCCGA 1803

Qy       42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
            ::::::||||||   :::|||:::|||:::|||   :::||||||||||||||||||||||
Db     1804 GAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAATGCCTATGATTTC 1863

Qy       62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
            ||||||:::   |||:::||||||:::|||||||||||||||:::|||||||||:::|||
Db     1864 ATCATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGT 1923

Qy       82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1924 GGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAACCCCAAGATCCTT 1983

Qy      102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
            :::||||||||||||||||||||||||||:::||||||::::::   ||||||||||||
Db     1984 TTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTGGTTCAGGCCGCT 2043
```

```
Qy     122 LeuGluLys 124
           |||:::|||
Db    2044 CTGGATAAG 2052

RESULT 14
AAF86127
ID   AAF86127 standard; cDNA; 4186 BP.
XX
AC   AAF86127;
XX
DT   25-JUN-2001  (first entry)
XX
DE   Cynomologous monkey P-glycoprotein cDNA.
XX
KW   Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   efflux pump; ss.
XX
OS   Macaca fascicularis.
XX
FH   Key             Location/Qualifiers
FT   CDS             100..3942
FT                   /*tag= a
FT                   /product= "PGP"
FT                   /note= "P-glycoprotein"
XX
PN   WO200123565-A1.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26592.
XX
PR   28-SEP-1999;   99US-0156921.
PR   12-OCT-1999;   99US-0158818.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-Crespi DT,  Crespi CL;
XX
DR   WPI; 2001-316136/33.
DR   P-PSDB; AAB81064.
XX
PT   Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein
PT   (PGP) and homologous PGP polypeptides are useful for predicting
PT   bioavailability of compound and increasing PGP transporter activity in
PT   cell  -
XX
PS   Example 1; Page 51-57; 84pp; English.
XX
CC   This invention relates to a polynucleotide sequence encoding a
CC   cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the
CC   PGP protein. PGP, also known as multidrug transporter, MDR1 is a member
CC   of the ABC transporter superfamily. The enzyme serves as an efflux pump
CC   exporting small molecules across the cell membrane. The invention
CC   includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence
CC   and protein, and also that of an allelic variant. The PGP polynucleotide
CC   sequence is useful for increasing PGP transporter activity in a cell.
CC   Antisense sequences of the cDNA are useful for inhibiting PGP transport
CC   activity in a mammalian cell. They may also be used for increasing the
CC   bioavailability of a drug. The present sequence represents cDNA encoding
CC   cynomologous monkey P-glycoprotein.
XX
SQ   Sequence 4186 BP; 1226 A; 799 C; 1039 G; 1122 T; 0 other;

Alignment Scores:
Pred. No.:              1.66e-51      Length:        4186
Score:                  465.00        Matches:       91
Percent Similarity:     88.62%        Conservative:  18
Best Local Similarity:  73.98%        Mismatches:    14
Query Match:            70.56%        Indels:        0
DB:                     22            Gaps:          0

US-09-873-409-3 (1-131) x AAF86127 (1-4186)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
           ||||||     |||||||||    :::|||||||||       |||:::   |||||||||
Db    1456 GTTGATGGACAGGATATTAGGACCATAAACGTAAGGTTTCTACGGGAAATCATCGGTGTG 1515

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           ||||||||||||||||||||||||   ||||||||||:::   ||||||:::|||||||||
Db    1516 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGTCGT 1575

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
           :::|||||||||   :::|||:::|||:::|||   :::|||||||||||||||||||||
Db    1576 GAAGATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1635

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           ||||||:::      |||   ||||||:::||||||||||||||||:::||||||||:::|||
Db    1636 ATCATGAAACTGCCTCAGAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGCTGAGT 1695

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1696 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 1755

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
           :::|||||||||||||||||||||||||||:::||||||::::::   ||||||   |||
Db    1756 CTGCTGGACGAGGCCACGTCAGCCTTGGACACAGAAAGTGAAGCAGTGGTTCAGGTGGCT 1815

Qy     122 LeuGluLys 124
           |||:::|||
Db    1816 CTGGATAAG 1824

RESULT 15
AAF86128
ID   AAF86128 standard; cDNA; 4195 BP.
XX
AC   AAF86128;
XX
DT   25-JUN-2001  (first entry)
XX
DE   Cynomologous monkey P-glycoprotein cDNA variant 1.
XX
KW   Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   efflux pump; ss.
XX
OS   Macaca fascicularis.
XX
FH   Key             Location/Qualifiers
FT   CDS             100..3951
FT                   /*tag= a
FT                   /product= "PGP"
FT                   /note= "P-glycoprotein"
FT   misc_feature    376..384
FT                   /*tag= b
FT                   /note= "Insertion of 9 nucleotides relative to PGP
FT                          allelic variant AAF86127"
XX
PN   WO200123565-A1.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26592.
XX
PR   28-SEP-1999;   99US-0156921.
PR   12-OCT-1999;   99US-0158818.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-Crespi DT,  Crespi CL;
XX
DR   WPI; 2001-316136/33.
DR   P-PSDB; AAB81065.
XX
PT   Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein
PT   (PGP) and homologous PGP polypeptides are useful for predicting
PT   bioavailability of compound and increasing PGP transporter activity in
```

```
PT   cell  -
XX
PS   Example 1; Page 59-65; 84pp; English.
XX
CC   This invention relates to a polynucleotide sequence encoding a
CC   cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the
CC   PGP protein. PGP, also known as multidrug transporter, MDR1 is a member
CC   of the ABC transporter superfamily. The enzyme serves as an efflux pump
CC   exporting small molecules across the cell membrane. The invention
CC   includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence
CC   and protein, and also that of an allelic variant. The PGP polynucleotide
CC   sequence is useful for increasing PGP transporter activity in a cell.
CC   Antisense sequences of the cDNA are useful for inhibiting PGP transport
CC   activity in a mammalian cell. They may also be used for increasing the
CC   bioavailability of a drug. The present sequence represents cDNA encoding
CC   cynomologous monkey P-glycoprotein. This sequence contains a 9 nucleotide
CC   insert compared to the PGP allelic variant given in AAF86127.
XX
SQ   Sequence 4195 BP; 1230 A; 801 C; 1039 G; 1125 T; 0 other;

Alignment Scores:
Pred. No.:              1.66e-51       Length:        4195
Score:                  465.00         Matches:       91
Percent Similarity:     88.62%         Conservative:  18
Best Local Similarity:  73.98%         Mismatches:    14
Query Match:            70.56%         Indels:        0
DB:                     22             Gaps:          0

US-09-873-409-3 (1-131) x AAF86128 (1-4195)

Qy        2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
            ||||||      |||||||||    :::|||||||||       |||:::    |||||||||
Db     1465 GTTGATGGACAGGATATTAGGACCATAAACGTAAGGTTTCTACGGGAAATCATCGGTGTG 1524

Qy       22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
            ||||||||||||||||||||||||   |||||||||:::    ||||||:::|||||||||
Db     1525 GTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGTCGT 1584

Qy       42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
            :::|||||||||   :::|||:::|||:::|||   :::||||||||||||||||||||||
Db     1585 GAAGATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1644

Qy       62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
            ||||||:::   |||   ||||||:::|||||||||||||||:::|||||||||:::|||
Db     1645 ATCATGAAACTGCCTCAGAAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGCTGAGT 1704

Qy       82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1705 GGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTC 1764

Qy      102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
            :::|||||||||||||||||||||||||||:::||||||::::::   ||||||   |||
Db     1765 CTGCTGGACGAGGCCACGTCAGCCTTGGACACAGAAAGTGAAGCAGTGGTTCAGGTGGCT 1824

Qy      122 LeuGluLys 124
            |||:::|||
Db     1825 CTGGATAAG 1833

Search completed: March 30, 2003, 03:08:18
Job time : 142.031 secs
```

```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 30, 2003, 01:24:25 ; Search time 1770.98 Seconds
                                           (without alignments)
                                           2152.747 Million cell updates/sec

Title:          US-09-873-409-3
Perfect score:  659
Sequence:       1 MVDENDIRALNVRHYRDHIG..........SESKSAVQAALEKDTPRYSF 131

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters:      4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115418_19228/app_query.fasta_1.7544
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_54887_@runat_27032003_115418_19228 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      GenEmbl:*
                1:  gb_ba:*
                2:  gb_htg:*
                3:  gb_in:*
                4:  gb_om:*
                5:  gb_ov:*
                6:  gb_pat:*
                7:  gb_ph:*
                8:  gb_pl:*
                9:  gb_pr:*
                10:  gb_ro:*
                11:  gb_sts:*
                12:  gb_sy:*
                13:  gb_un:*
                14:  gb_vi:*
                15:  em_ba:*
                16:  em_fun:*
                17:  em_hum:*
                18:  em_in:*
                19:  em_mu:*
                20:  em_om:*
                21:  em_or:*
                22:  em_ov:*
                23:  em_pat:*
                24:  em_ph:*
                25:  em_pl:*
                26:  em_ro:*
                27:  em_sts:*
                28:  em_un:*
                29:  em_vi:*
                30:  em_htg_hum:*
                31:  em_htg_inv:*
                32:  em_htg_other:*
                33:  em_htg_mus:*
                34:  em_htg_pln:*
                35:  em_htg_rod:*
                36:  em_htg_mam:*
                37:  em_htg_vrt:*
                38:  em_sy:*
                39:  em_htgo_hum:*
                40:  em_htgo_mus:*
                41:  em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	659	100.0	1175	6	AX339029	AX339029 Sequence
2	659	100.0	1940	6	AX339034	AX339034 Sequence
3	659	100.0	2021	6	AX339033	AX339033 Sequence
4	619	93.9	2856	6	AX339028	AX339028 Sequence
5	619	93.9	3177	6	AX339030	AX339030 Sequence
6	619	93.9	3621	6	AX339032	AX339032 Sequence
7	619	93.9	3699	6	AX478104	AX478104 Sequence
8	619	93.9	3702	6	AX339031	AX339031 Sequence
9	474	71.9	3489	4	AB029153	AB029153 Felis cat
10	472	71.6	3888	10	AY082609	AY082609 Rattus no
11	472	71.6	4233	6	AR123273	AR123273 Sequence
12	469	71.2	2752	5	AF099732	AF099732 Fundulus
13	469	71.2	3934	4	CFA419568	AJ419568 Canis fam
14	469	71.2	4018	5	GGA9799	AJ009799 Gallus ga
15	469	71.2	4045	12	AF269224	AF269224 Synthetic
16	469	71.2	4279	6	AX105057	AX105057 Sequence
17	469	71.2	4279	6	AX105078	AX105078 Sequence
18	469	71.2	4279	6	AX105080	AX105080 Sequence
19	469	71.2	4279	6	AX105082	AX105082 Sequence
20	468	71.0	4189	6	AX322791	AX322791 Sequence
21	468	71.0	4298	10	MUSMDR	M14757 Mouse multi
22	468	71.0	4356	10	MUSMDRAA	M30697 Mouse multi
23	467	70.9	3858	4	OAU78609	U78609 Ovis aries
24	467	70.9	3987	10	CRUPGP1185	M59254 Chinese ham
25	467	70.9	4254	6	AX401748	AX401748 Sequence
26	467	70.9	4254	10	RATMDRM	M81855 Rat mdr mRN
27	467	70.9	4280	10	CRUPGPII	M60041 C.griseus P
28	467	70.9	4296	10	CRUPGPI	M60040 C.griseus P
29	467	70.9	4304	10	CRUPGP1165	M59253 Chinese ham
30	467	70.9	4323	10	AF286167	AF286167 Rattus no
31	467	70.9	4927	10	AF257746	AF257746 Rattus no
32	465	70.6	2005	4	AB066299	AB066299 Canis fam
33	465	70.6	4186	6	AX108654	AX108654 Sequence
34	465	70.6	4195	6	AX108656	AX108656 Sequence
35	465	70.6	4317	4	AF045016	AF045016 Canis fam
36	465	70.6	4317	6	AX105059	AX105059 Sequence
37	465	70.6	4390	5	XLU17608	U17608 Xenopus lae
38	462	70.1	3682	10	MUSPGLY1A	M24417 Mouse phosp
39	462	70.1	4788	6	AX322793	AX322793 Sequence
40	462	70.1	4924	10	MUSMDR1A	M33581 Mouse P-gly
41	461	70.0	3840	6	AX481416	AX481416 Sequence
42	461	70.0	3860	6	AX322787	AX322787 Sequence
43	461	70.0	3860	6	AX322789	AX322789 Sequence
44	461	70.0	3988	6	AX024454	AX024454 Sequence
45	461	70.0	4192	9	AF016535	AF016535 Homo sapi

ALIGNMENTS

RESULT 1

```
AX339029
LOCUS       AX339029              1175 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 11 from Patent WO0194400.
ACCESSION   AX339029
VERSION     AX339029.1  GI:18129121
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 11 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..1175
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      337 a    238 c    258 g    342 t
ORIGIN

Alignment Scores:
Pred. No.:              4.24e-60      Length:        1175
Score:                  659.00        Matches:       131
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     6             Gaps:          0

US-09-873-409-3 (1-131) x AX339029 (1-1175)

Qy       1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     418 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 477

Qy      21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     478 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 537

Qy      41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     538 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 597

Qy      61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     598 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 657

Qy      81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     658 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 717

Qy     101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     718 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 777

Qy     121 AlaLeuGluLysAspThrProArgTyrSerPhe 131
           |||||||||||||||||||||||||||||||||
Db     778 GCACTGGAGAAGGATACCCCCAGGTATTCATTT 810

RESULT 2
AX339034
LOCUS       AX339034              1940 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 16 from Patent WO0194400.
ACCESSION   AX339034
VERSION     AX339034.1  GI:18129126
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 16 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..1940
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      530 a    394 c    448 g    568 t
ORIGIN

Alignment Scores:
Pred. No.:              7.66e-60      Length:        1940
Score:                  659.00        Matches:       131
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     6             Gaps:          0

US-09-873-409-3 (1-131) x AX339034 (1-1940)

Qy       1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1183 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 1242

Qy      21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1243 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 1302

Qy      41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1303 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 1362

Qy      61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1363 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 1422

Qy      81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1423 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 1482

Qy     101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1483 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 1542

Qy     121 AlaLeuGluLysAspThrProArgTyrSerPhe 131
           |||||||||||||||||||||||||||||||||
Db    1543 GCACTGGAGAAGGATACCCCCAGGTATTCATTT 1575

RESULT 3
AX339033
LOCUS       AX339033              2021 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 15 from Patent WO0194400.
ACCESSION   AX339033
VERSION     AX339033.1  GI:18129125
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 15 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..2021
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      547 a    406 c    462 g    605 t      1 others
ORIGIN
```

```
Alignment Scores:
Pred. No.:              8.03e-60      Length:        2021
Score:                  659.00        Matches:       131
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            100.00%       Indels:        0
DB:                     6             Gaps:          0

US-09-873-409-3 (1-131) x AX339033 (1-2021)

Qy       1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1264 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 1323

Qy      21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1324 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 1383

Qy      41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1384 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 1443

Qy      61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1444 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 1503

Qy      81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1504 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 1563

Qy     101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1564 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 1623

Qy     121 AlaLeuGluLysAspThrProArgTyrSerPhe 131
             |||||||||||||||||||||||||||||||||
Db    1624 GCACTGGAGAAGGATACCCCCAGGTATTCATTT 1656

RESULT 4
AX339028
LOCUS       AX339028                2856 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 10 from Patent WO0194400.
ACCESSION   AX339028
VERSION     AX339028.1  GI:18129120
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 10 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..2856
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      888 a    540 c    652 g    776 t
ORIGIN

Alignment Scores:
Pred. No.:              1.9e-55       Length:        2856
Score:                  619.00        Matches:       124
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            93.93%        Indels:        0
DB:                     6             Gaps:          0

US-09-873-409-3 (1-131) x AX339028 (1-2856)

Qy       1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     418 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 477

Qy      21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     478 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 537

Qy      41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     538 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 597

Qy      61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     598 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 657

Qy      81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     658 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 717

Qy     101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     718 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 777

Qy     121 AlaLeuGluLys 124
             |||||||||||
Db     778 GCACTGGAGAAG 789

RESULT 5
AX339030
LOCUS       AX339030                3177 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 12 from Patent WO0194400.
ACCESSION   AX339030
VERSION     AX339030.1  GI:18129122
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 12 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..3177
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      970 a    601 c    735 g    870 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              2.15e-55      Length:        3177
Score:                  619.00        Matches:       124
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            93.93%        Indels:        0
DB:                     6             Gaps:          0

US-09-873-409-3 (1-131) x AX339030 (1-3177)

Qy       1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     739 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 798

Qy      21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     799 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 858

Qy      41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     859 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 918
```

```
Qy       61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      919 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 978

Qy       81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      979 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 1038

Qy      101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1039 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 1098

Qy      121 AlaLeuGluLys 124
            ||||||||||||
Db     1099 GCACTGGAGAAG 1110
```

```
RESULT 6
AX339032
LOCUS       AX339032                3621 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 14 from Patent WO0194400.
ACCESSION   AX339032
VERSION     AX339032.1  GI:18129124
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 14 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..3621
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1081 a    696 c    842 g   1002 t
ORIGIN

Alignment Scores:
Pred. No.:              2.51e-55       Length:          3621
Score:                  619.00         Matches:         124
Percent Similarity:     100.00%        Conservative:    0
Best Local Similarity:  100.00%        Mismatches:      0
Query Match:            93.93%         Indels:          0
DB:                     6              Gaps:            0

US-09-873-409-3 (1-131) x AX339032 (1-3621)

Qy        1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1183 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 1242

Qy       21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1243 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 1302

Qy       41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1303 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 1362

Qy       61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1363 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 1422

Qy       81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1423 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 1482

Qy      101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1483 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 1542

Qy      121 AlaLeuGluLys 124
            ||||||||||||
Db     1543 GCACTGGAGAAG 1554
```

```
RESULT 7
AX478104
LOCUS       AX478104                3699 bp    DNA     linear   PAT 12-AUG-2002
DEFINITION  Sequence 34 from Patent WO0240541.
ACCESSION   AX478104
VERSION     AX478104.1  GI:22217064
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Tang,Y.T., Yue,H., Nguyen,D.B., Hafalia,A.J., Elliott,V.S., Lu,Y.,
            Walia,N.K., Yao,M.G., Baughn,M.R., Gandhi,A.R., Ding,L.,
            Sanjanwala,M., Ramkumar,J., Arvizu,C., Gietzen,K.J., Lal,P.G.,
            Azimzai,Y., Khan,F.A., Thangavelu,K., Thornton,M., Lu,D.A.,
            Tribouley,C.M., Warren,B.A., Ison,C.H., Das,D., Raumann,B.E.,
            Policky,J.L. and Kearney,L.
  TITLE     Transporters and ion channels
  JOURNAL   Patent: WO 0240541-A 34 23-MAY-2002;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..3699
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 7472030CB1"
BASE COUNT     1116 a    707 c    860 g   1016 t
ORIGIN

Alignment Scores:
Pred. No.:              2.58e-55       Length:          3699
Score:                  619.00         Matches:         124
Percent Similarity:     100.00%        Conservative:    0
Best Local Similarity:  100.00%        Mismatches:      0
Query Match:            93.93%         Indels:          0
DB:                     6              Gaps:            0

US-09-873-409-3 (1-131) x AX478104 (1-3699)

Qy        1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1339 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 1398

Qy       21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1399 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 1458

Qy       41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1459 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 1518

Qy       61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1519 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 1578

Qy       81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1579 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 1638

Qy      101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1639 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 1698

Qy      121 AlaLeuGluLys 124
            ||||||||||||
Db     1699 GCACTGGAGAAG 1710
```

```
RESULT 8
AX339031
LOCUS       AX339031                3702 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 13 from Patent WO0194400.
ACCESSION   AX339031
VERSION     AX339031.1  GI:18129123
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 13 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..3702
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1098 a    708 c    856 g   1039 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              2.58e-55      Length:        3702
Score:                  619.00        Matches:       124
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            93.93%        Indels:        0
DB:                     6             Gaps:          0

US-09-873-409-3 (1-131) x AX339031 (1-3702)

Qy      1 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1264 ATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGA 1323

Qy     21 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1324 GTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGA 1383

Qy     41 ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAsp 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1384 CGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGAT 1443

Qy     61 PheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMet 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1444 TTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATG 1503

Qy     81 SerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIle 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1504 AGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATT 1563

Qy    101 LeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAla 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1564 CTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCT 1623

Qy    121 AlaLeuGluLys 124
          ||||||||||||
Db   1624 GCACTGGAGAAG 1635

RESULT 9
AB029153
LOCUS       AB029153                3489 bp    mRNA    linear   MAM 28-MAR-2002
DEFINITION  Felis catus multi-drug resistance related mRNA, partial cds.
ACCESSION   AB029153
VERSION     AB029153.1  GI:6472652
KEYWORDS    .
SOURCE      Felis catus feline lymphoma cell line
            cell_line:adriamycin-resistant subline of FT-1 cDNA to mRNA.
  ORGANISM  Felis catus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
REFERENCE   1
  AUTHORS   Okai,Y., Nakamura,N., Matsushiro,H., Kato,H., Setoguchi,A.,
            Yazawa,M., Okuda,M., Watari,T., Hasegawa,A. and Tsujimoto,H.
  TITLE     Molecular analysis of multidrug resistance in feline lymphoma cells
  JOURNAL   Am. J. Vet. Res. 61 (9), 1122-1127 (2000)
  MEDLINE   20431258
REFERENCE   2  (bases 1 to 3489)
  AUTHORS   Tsujimoto,H. and Okai,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-JUN-1999) Hajime Tsujimoto, The University of Tokyo,
            Department of Veterinary Internal Medicine; The University of
            Tokyo, Bunkyo-ku, Tokyo 113-8657, Japan
            (E-mail:atsuji@hongo.ecc.u-tokyo.ac.jp, Tel:+81-3-5841-8004,
            Fax:+81-3-5841-8178)
FEATURES             Location/Qualifiers
     source          1..3489
                     /organism="Felis catus"
                     /db_xref="taxon:9685"
                     /cell_line="adriamycin-resistant subline of FT-1"
                     /cell_type="feline lymphoma cell line"
     CDS             <1..>3489
                     /note="multi-drug resistance related mRNA"
                     /codon_start=1
                     /protein_id="BAA87071.1"
                     /db_xref="GI:6472653"
                     /translation="HGAALPLMMLVFGDMTDSFANAGISRNLTTLNITGESIVNDSYF
                     INRLEEEMTTYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQILKIRKQFFHAIMRQEVG
                     WFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSMATFFIGFIVGFTRGWKLTLVIL
                     AISPVLGLSAAIWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFGGQKKELERY
                     NKNLEEAKRIGIKKAITANISIGIAFLLIYASYALAFWYGTSLVLSHEYSIGQVLTVF
                     FSVLIGAFSVGQASPSIEAFANARGAAYEIFKIIDNKPSIDSYSKNGHKPDNIKGNLE
                     FKNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQRLYDPTDGMV
                     SIDGQDIRTINVRYLREIIGVVSQEPVLFATTIAENIRYGRENVTMEEIEKAVKEANA
                     YDFIMKLPNKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDSESEA
                     VVQVALDKARKGRTTIVVAHRLSTIRNADVIAGFDDGVIVEKGNHDELMKEEGIYFKL
                     VTMQTRGNEIELENAVYESISEIDALEMSPKDSGSSLIRRRSTRKSIHASQGQDRKLG
                     TKENLDERVPPVSFWRILKLNITEWPYFVVGIFCAIINGGLQPAFSVILSRIIGVFTR
                     VEDPETKRQNSNIFSLLFLVLGIISSITFFLQGFTFGKAGEILTKRLRYMVFRSMLRQ
                     DVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAIITQNIANLGTGIIISLIYGWQL
                     TLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKVATEAIENFRTVVSLTREQK
                     FEYMYAQSLQVPYRNSLRKAHIFGISFSITQAMMYFSYAGCFRFGAYLVAHEFMDFQD
                     VLLVFSAIVFGAMAVGQVSSFAPDYAKAKVSAAHVIMIIEKIPLIDSYSTEGLMPNTL
                     EGNVTFNEVMFNYPTRPDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERFYDP
                     MAGTVLIDGKEIKHLNVQWLRAHMGIVSQEPILFDCSIGENIAYGDNSRVVSQEEIVR
                     AAKEANIHPFIETLPDKYNTRVGDKGTQLSGGQKQRIAIARALVRQPRILLLDEATSA
                     LDTESEKVVQEALDKAR"
BASE COUNT     1015 a    716 c    847 g    911 t
ORIGIN

Alignment Scores:
Pred. No.:              3.95e-40      Length:        3489
Score:                  474.00        Matches:       92
Percent Similarity:     90.24%        Conservative:  19
Best Local Similarity:  74.80%        Mismatches:    12
Query Match:            71.93%        Indels:        0
DB:                     4             Gaps:          0

US-09-873-409-3 (1-131) x AB029153 (1-3489)

Qy      2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
          :::|||      |||||||||   :::|||||||||:::   |||:::   |||||||||
Db   1180 ATTGATGGACAGGACATTAGGACCATAAATGTGAGATATCTTCGGGAAATCATTGGTGTG 1239

Qy     22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
          ||||||||||||||||||||||||   |||||||||:::   ||||||:::|||||||||
Db   1240 GTGAGTCAGGAACCTGTGTTGTTTGCCACTACAATAGCTGAAAACATTCGCTATGGCCGT 1299

Qy     42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
          ::::::||||||   ||||||:::|||:::|||   :::|||||||||||||||||||||
Db   1300 GAAAATGTCACCATGGAGGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTT 1359

Qy     62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
```

```
         ||||||:::   ||||||||||||:::||||||||||||||:::|||||||||:::|||
Db  1360 ATCATGAAACTCCCTAATAAATTTGATACCTTGGTTGGAGAGCGGGGGGCCCAACTGAGC 1419

Qy    82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1420 GGTGGACAGAAACAGAGAATCGCCATCGCTCGGGCCCTGGTTCGCAACCCCAAGATCCTC 1479

Qy   102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
         :::|||||||||||||||||||||||||||||||||||::::::   ||||||   |||
Db  1480 CTGCTGGATGAGGCCACATCAGCCTTGGACAGTGAAAGTGAAGCAGTGGTTCAGGTGGCC 1539

Qy   122 LeuGluLys 124
         |||:::|||
Db  1540 CTGGATAAG 1548

RESULT 10
AY082609
LOCUS       AY082609                3888 bp    mRNA    linear   ROD 26-MAR-2002
DEFINITION  Rattus norvegicus ATP-binding cassette protein B1b mRNA, complete
            cds.
ACCESSION   AY082609
VERSION     AY082609.1  GI:19743729
KEYWORDS    .
SOURCE      Rattus norvegicus.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 3888)
  AUTHORS   Yabuuchi,H. and Ishikawa,T.
  TITLE     Direct Submission
  JOURNAL   Submitted (06-MAR-2002) Biomolecular Engineering, Tokyo Institute
            of Technology, 4259 Nagatsuta, Midori-ku, Yokohama, Kanagawa
            226-8501, Japan
FEATURES             Location/Qualifiers
     source          1..3888
                     /organism="Rattus norvegicus"
                     /strain="Sprague-Dawley"
                     /db_xref="taxon:10116"
                     /tissue_type="liver"
     CDS             36..3863
                     /note="transporter; ABCB1b"
                     /codon_start=1
                     /product="ATP-binding cassette protein B1b"
                     /protein_id="AAL92458.1"
                     /db_xref="GI:19743730"
                     /translation="MEFEEGLNGTADKNFSKMGKKSKKEKEKKPAVGIFGMFRYADWL
                     DKLCMALGTLAAIIHGTLLPLLMLVFGYMTDSFTQAETRILPSVTNQSEINSTQTVSD
                     SSLEEDMAMYAYYYTGIGAGVLIVAYIQVSLWCLAAGRQIHKIRQKFFHAIMNQEIGW
                     FDVNDAGELNTRLTDDVSKINDGIGDKLGMFFQSITTFSAGFIIGFISGWKLTLVILA
                     VSPLIGLSSAMWAKVLTSFTNKELQAYAKAGAVAEEVLAAIRTVIAFGGQKKELERYN
                     KNLEEAKRVGIKKAITANISIGIAYLLVYASYALAFWYGTSLVLSNEYSIGQVLTVFF
                     SILLGTFSIGHLAPNIEAFANARGAAYEIFKIIDNEPSIDSFSTKGHKPDSIMGNLEF
                     KNVYFNYPSRSEVKILKGLNLKVKSGQTVALVGNSGCGKSTTVQLLQRLYDPIEGEVS
                     IDGQDIRTINVRYLREIIGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKEANAY
                     DFIMKLPHKFNTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTESEAV
                     VQAALDKAREGRTTIVIAHRLSTVRNADVIAGFDGGVIVEQGNHEELMKEKGIYFKLV
                     MTQTRGNEIEPGNNAYESQSDTGASELTSEKSKSPLIRRSIRRSIHRRQDQERRLSSK
                     EDVDEDVPMVSFWQILKLNISEWPYLVVGVLCAVINGCIQPVFAIVFSKIVGVFSRDD
                     DHETKQRNCNLFSLLFLVMGMISFVTYFFQGFAFGKAGEILTKRLRYMVFKSMLRQDI
                     SWFDDHKNTTGSLTTRLASDASNVKGAMGSRLAVVTQNVANLGTGIILSLVYGWQLTL
                     LLVVIIPLIVLGGIIEMKLLSGQALKDKKELEISGKIATEAIENFRTVVSLTREQKFE
                     TMYAQSLQIPYRNALKKAHVFGITFAFTQAMIYFSYAACFRFGAYLVARELMTFENVM
                     LVFSAVVFGAMAAGNTSSFAPDYAKAKVSASHIIRIIEKIPEIDSYSTEGLKPNWLEG
                     NVKFNGVMFNYPTRPNIPVLQGLSFEVKKGQTLALVGSSGCGKSTVVQLLERFYNPMA
                     GTVFLDGKEIKQLNVQWLRAHLGIVSQEPILFDCSITENIAYGDNSRVVSHEEIVRAA
                     REANIHQFIDSLPEKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEATSALD
                     TESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVIQNGQVKEHGTHQQLLAQKG
                     IYFSMVQAGAKRS"
BASE COUNT     1071 a    877 c   1003 g    937 t
ORIGIN

Alignment Scores:
Pred. No.:               7.27e-40       Length:          3888
Score:                   472.00         Matches:         91
Percent Similarity:      91.06%         Conservative:    21
Best Local Similarity:   73.98%         Mismatches:      11
Query Match:             71.62%         Indels:          0
DB:                      10             Gaps:            0

US-09-873-409-3 (1-131) x AY082609 (1-3888)

Qy     2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
         :::|||      |||||||||   :::|||||||||:::   |||:::   |||||||||
Db  1386 ATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTG 1445

Qy    22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
         ||||||||||||||||||||||||      |||||||||:::   ||||||:::||||||
Db  1446 GTGAGTCAGGAACCCGTGCTGTTTGCCACCACGATTGCCGAAAACATTCGCTATGGCCGA 1505

Qy    42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
         ::::::||||||   :::|||:::|||:::|||   :::|||||||||||||||||||||
Db  1506 GAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTC 1565

Qy    62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
         ||||||:::   |||:::||||||||||||||||||||||||:::|||||||||:::|||
Db  1566 ATCATGAAACTGCCCCACAAATTTAACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGT 1625

Qy    82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1626 GGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAACCCCAAGATCCTT 1685

Qy   102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
         :::||||||||||||||||||||||||||||:::||||||::::::   ||||||||||||
Db  1686 TTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTGGTTCAGGCCGCT 1745

Qy   122 LeuGluLys 124
         |||:::|||
Db  1746 CTGGATAAG 1754

RESULT 11
AR123273
LOCUS       AR123273                4233 bp    DNA     linear   PAT 16-MAY-2001
DEFINITION  Sequence 1 from patent US 6169166.
ACCESSION   AR123273
VERSION     AR123273.1  GI:14108239
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 4233)
  AUTHORS   Brun,K.A., Chenery,R.J., Ellens,H., Feild,J.A. and Yue,L.
  TITLE     Polynucleotide and polypeptide sequences encoding rat mdr1b2 and
            screening methods thereof
  JOURNAL   Patent: US 6169166-A 1 02-JAN-2001;
FEATURES             Location/Qualifiers
     source          1..4233
                     /organism="unknown"
BASE COUNT     1210 a    920 c   1048 g   1055 t
ORIGIN

Alignment Scores:
Pred. No.:               8.04e-40       Length:          4233
Score:                   472.00         Matches:         91
Percent Similarity:      91.06%         Conservative:    21
Best Local Similarity:   73.98%         Mismatches:      11
Query Match:             71.62%         Indels:          0
DB:                      6              Gaps:            0

US-09-873-409-3 (1-131) x AR123273 (1-4233)

Qy     2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
         :::|||      |||||||||   :::|||||||||:::   |||:::   |||||||||
Db  1376 ATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTG 1435

Qy    22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
```

```
                 ||||||||||||||||||||||   |||||||||:::  ||||||:::||||||||
Db    1436 GTGAGTCAGGAACCCGTGCTGTTTGCCACCACGATTGCCGAAAACATTCGCTATGGCCGA 1495

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
           ::::::||||||    :::|||:::|||:::|||    :::||||||||||||||||||||
Db    1496 GAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTC 1555

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           ||||||:::   |||:::||||||||||||||||||||||||:::|||||||||:::|||
Db    1556 ATCATGAAACTGCCCCACAAATTTAACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGT 1615

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1616 GGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAACCCCAAGATCCTT 1675

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
           :::||||||||||||||||||||||||||||:::||||||::::::   |||||||||||
Db    1676 TTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTGGTTCAGGCCGCT 1735

Qy     122 LeuGluLys 124
           |||:::|||
Db    1736 CTGGATAAG 1744
```

RESULT 12
AF099732

```
LOCUS       AF099732              2752 bp    mRNA    linear   VRT 08-APR-1999
DEFINITION  Fundulus heteroclitus multidrug resistance transporter homolog
            (mdr) mRNA, partial cds.
ACCESSION   AF099732
VERSION     AF099732.1  GI:4574223
KEYWORDS    .
SOURCE      Fundulus heteroclitus.
  ORGANISM  Fundulus heteroclitus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
            Cyprinodontiformes; Fundulidae; Fundulus.
REFERENCE   1  (bases 1 to 2752)
  AUTHORS   Cooper,P.S., Van Veld,P.A. and Reece,K.S.
  TITLE     P-glycoprotein related sequences from the mummichog (Fundulus
            heteroclitus)
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 2752)
  AUTHORS   Cooper,P.S., Van Veld,P.A. and Reece,K.S.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-OCT-1998) School of Marine Science, Department of
            Environmental Sciences, Virginia Institute of Marine Science,
            College of William and Mary, Gloucester Point, VA 23062, USA
FEATURES             Location/Qualifiers
     source          1..2752
                     /organism="Fundulus heteroclitus"
                     /db_xref="taxon:8078"
                     /tissue_type="liver; intestine"
     gene            <1..2752
                     /gene="mdr"
     CDS             <1..2556
                     /gene="mdr"
                     /codon_start=1
                     /product="multidrug resistance transporter homolog"
                     /protein_id="AAD23956.1"
                     /db_xref="GI:4574224"
                     /translation="STTIQLLERFYDPQEGSVSIDGHDIRSLNVRYLRGMIGVVSQEP
                     ILFATTIAENIRYGRPDVTEEEIEKAAKEANAYDFIMNLPDKFETLVGDRGTQMSGGQ
                     KQRIAIARALVRKPKILLLDEATSALDAESETIVQAALDKVRQGRTTLIVAHCLSTIR
                     NADVIAGLEKGKVVELGTHSELMEKKGVYHTLVTMQTFQKADEGEDEDNLSAGEKSPI
                     HNNVIESPLLRRKSTRGSSFAASIGEKGDKKQEKEDEDKTEEDEDFPMVSIFKVLRLN
                     ASEWPYILVGLICATINGAIQPLFAVLFSKIITVFAEPDQTIIRQRANFFSLMFVVIG
                     VVCFFTMFLQGFCFGKSGEVLTLKLRLGAFKSMLRQDLGWFDSPKNSVGALTTRLATD
                     AAQVQGASGVRLATFAQNIANLGTGVILAFVYGWELTLLILAVVPVIALAGAVQMKML
                     TGHAAEDKKELEKAGKIATEAIENIRTVASLTREPKFESLYEENLVVPYKNSQKKAHV
                     YGFTFSFSQAMIYFAYAACFRFGAWLIVEGRMDVEAVFLVISAVLFGAMAVGEANSFA
                     PNYAKAKMSASHLMMLLNKEPEIDNLSERGESPDMFDGNVSFEDVKFNYPSRPDVPIL
                     RGLNLRVKKGETLALVGSSGCGKSTTIQLLERFYDPRDGRVVMDSIDVKRLNIHWLRS
                     QIGIVSQEPVLFDCALAENIAYGDNSRSVTMEEIEAAAKAANIHNFINELPQKYNTQA
                     GDKGTQLSGGQKQRIAIARAILRNPKVLLLDEATSALDTESEKVVQDALDQASKGRTC
                     IVVAHRLSTIRNADRIAVFQGGVVVEQGTHQQLLAKKGVYHMLVTTQLGHGTE"
BASE COUNT      661 a    718 c    791 g    582 t
ORIGIN
```

Alignment Scores:

Pred. No.:	1e-39	Length:	2752
Score:	469.00	Matches:	91
Percent Similarity:	89.43%	Conservative:	19
Best Local Similarity:	73.98%	Mismatches:	13
Query Match:	71.17%	Indels:	0
DB:	5	Gaps:	0

US-09-873-409-3 (1-131) x AF099732 (1-2752)

```
Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
           :::|||   :::||||||||:::|||||||||||||:::   |||      |||||||||
Db      58 ATCGATGGGCACGACATCCGCTCTCTTAATGTGCGCTACCTGAGGGGAATGATCGGCGTG 117

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           |||||||||||||||:::||||||   ||||||||||:::   ||||||:::|||||||||
Db     118 GTGAGCCAGGAGCCCATCCTCTTCGCCACCACCATCGCTGAGAACATCCGCTACGGCCGA 177

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
              |||||||||:::||||||:::|||:::||||||:::||||||||||||||||||||
Db     178 CCCGACGTGACGGAGGAGGAGATCGAAAAAGCTGCCAAGGAAGCCAATGCTTACGACTTC 237

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           ||||||      |||:::||||||   ||||||||||||::::::|||   |||||||||
Db     238 ATCATGAACCTTCCTGATAAGTTTGAGACACTAGTCGGCGACCGAGGAACCCAGATGAGC 297

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           |||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db     298 GGAGGACAGAAGCAGAGGATTGCGATCGCTCGAGCTTTGGTCCGAAAACCCAAAATCCTG 357

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
           :::||||||||||||||||||||||||||||:::||||||::::::   |||||||||||
Db     358 CTGTTGGACGAAGCCACGTCTGCCCTGGATGCTGAGAGTGAGACTATCGTTCAAGCTGCA 417

Qy     122 LeuGluLys 124
           |||:::|||
Db     418 CTCGATAAG 426
```

RESULT 13
CFA419568

```
LOCUS       CFA419568             3934 bp    mRNA    linear   MAM 06-DEC-2001
DEFINITION  Canis familiaris mRNA for abortive P-glycoprotein (p-gp gene).
ACCESSION   AJ419568
VERSION     AJ419568.1  GI:17385398
KEYWORDS    P-glycoprotein; p-gp gene.
SOURCE      dog.
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1
  AUTHORS   Roulet,A.
  TITLE     Characterization of a MDR1a P-gp deficient dogin relation to
            ivermectin sensitivity
  JOURNAL   Thesis (2002) Department of Biological Sciences, Universite de
            Toulouse (UPS III), Toulouse, France
REFERENCE   2  (bases 1 to 3934)
  AUTHORS   Roulet,A.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-NOV-2001) Roulet A., Pharmacology, Inra, 180 Chemin
            de Tournefeuille, 31931 Toulouse cedex, FRANCE
FEATURES             Location/Qualifiers
     source          1..3934
                     /organism="Canis familiaris"
                     /db_xref="taxon:9615"
                     /tissue_type="liver"
                     /note="ivermectin sensitive collies"
     gene            7..282
```

```
     CDS             /gene="p-gp"
                     7..282
                     /gene="p-gp"
                     /note="abortive product"
                     /codon_start=1
                     /evidence=experimental
                     /product="P-glycoprotein"
                     /protein_id="CAD12806.1"
                     /db_xref="GI:17385399"
                     /translation="MDPEGGRRGSAEKNFWKMGKKSKKNEKKGKKPTVSTFAMFRYSN
                     WLDRLYMLVGTMAAIIHGAALPLMMLVFGNMTALQMQEFQETKLFQL"
BASE COUNT     1144 a    815 c    964 g   1011 t
ORIGIN

Alignment Scores:
Pred. No.:              1.52e-39        Length:         3934
Score:                  469.00          Matches:        90
Percent Similarity:     89.43%          Conservative:   20
Best Local Similarity:  73.17%          Mismatches:     13
Query Match:            71.17%          Indels:         0
DB:                     4               Gaps:           0

US-09-873-409-3 (1-131) x CFA419568 (1-3934)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
           :::|||      ||||||||   :::|||||||||||   |||:::       ||||||
Db    1365 ATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTG 1424

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           ||||||||||||||||||||||||   ||||||||||:::   ||||||:::|||||||||
Db    1425 GTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGC 1484

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
           ::::::||||||   :::|||:::|||:::|||   :::||||||||||||||||||||||
Db    1485 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTT 1544

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           ||||||:::   ||||||||||||:::||||||||||||||||:::|||||||||:::|||
Db    1545 ATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGT 1604

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1605 GGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTT 1664

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
           :::|||||||||||||||||||||||||||:::||||||::::::   ||||||   |||
Db    1665 CTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCC 1724

Qy     122 LeuGluLys 124
           |||:::|||
Db    1725 CTGGATAAG 1733

RESULT 14
GGA9799
LOCUS       GGA9799                 4018 bp    mRNA    linear   VRT 03-MAR-1999
DEFINITION  Gallus gallus mRNA for ABC transporter protein.
ACCESSION   AJ009799
VERSION     AJ009799.1  GI:3355756
KEYWORDS    ABC transporter protein; cmdr1 gene; P-glycoprotein.
SOURCE      Gallus gallus.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 4018)
  AUTHORS   Edelmann,H.M.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JUL-1998) Edelmann H.M.L., Institute of Molecular
            Genetics, Biocenter and University of Vienna, Dr. Bohr-Gasse 9/2
            A-1030 Vienna, A-1030 Vienna, AUSTRIA
REFERENCE   2  (bases 1 to 4018)
  AUTHORS   Edelmann,H.M., Duchek,P., Rosenthal,F.E., Foger,N., Glackin,C.,
            Kane,S.E. and Kuchler,K.
  TITLE     Cmdr1, a chicken P-glycoprotein, confers multidrug resistance and
            interacts with estradiol
  JOURNAL   Biol. Chem. 380 (2), 231-241 (1999)
  MEDLINE   99209805
   PUBMED   10195430
FEATURES             Location/Qualifiers
     source          1..4018
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /tissue_type="intestine"
     gene            109..3975
                     /gene="cmdr1"
     CDS             109..3975
                     /gene="cmdr1"
                     /note="ABC transporter protein"
                     /codon_start=1
                     /product="P-glycoprotein"
                     /protein_id="CAA08835.1"
                     /db_xref="GI:3355757"
                     /db_xref="SPTREMBL:O93437"
                     /translation="MHSEDEKQKHTVDGNTYEIANSNQDQDPEDEKKGKKKKGKKPQM
                     VSPLALFRYSSCTDKLLMIFGSLLAIAHGTSLPIAMIIFGDMTDSFVTSGMTNITGNS
                     SGLNSSADVFNKLEEEMTRYAYYYSAIAAAVLVAAYIQTSFWTLAAGRQVKKIREKFF
                     HAIMRQEIGWFDVNDAGELNTRLIDDVSKINEGIGDKIGFLIQSETTFLTGFIVGFIR
                     GWKLTLVILAVSPVLGLSAALWAKILTAFTDKEQAAYAKAGAVAEEVLSAVRTVIAFG
                     GQEKEIKRYHKNLEDAKRIGIRKAITSNISMGAAFLLIYASYALAFWYGTTLILANEY
                     SIGNVLTVFFSVLIGAFSIGQTAPSIEAFANARGAAYAIFNIIDNEPEIDSYSDAGHK
                     PDHIKGNLEFQNVFFNYPSRPDVEILKGLNLKVNCGQTVALVGGSGCGKSTTVQLIQR
                     FYDPKEGTITIDGQDLKSLNVRYLREIIGVVNQEPVLFATTIAENIRYGREDVTMEEI
                     ERATKEANAYDFIMKLPKKFETVVGERGAQMSGGQKQRIAIARALVHNPKILLLDEAT
                     SALDTESESVVQAALDKAREGRTTVVVAHRLSTVRNADLIAVFESGVITEQGNHSQLI
                     EKKGIYYKLVNMQTIETEDPSSEKSENAVSVKRSGSQSNLDESLKKELRRGSTRRSMK
                     KPGEPNDTDEKGSSPDEELPPVSFLKLMKLNKNEWPYFVAGTFCAIVNGALQPAFSVI
                     FSEIIGIFSETDQKVLREKSNLYSLLFLALGIISFFTFFVQGFAFGKAGEILTMKLRF
                     MAFKAMLRQDMAWFDDPKNSTGALTTRLANDASQVKGATGVRLALIAQNIANLGTGII
                     ISLVYGWQLTLLLLAVVPIIAVAGMIEMKMLAGHAKKDKIELEAAGKIATEAIENIRT
                     VASLTREKRFELMYGEHLLVPYRNSVKKAHIFGFCFSLSQAMMFFTYAGCFRFGAYLV
                     VNGHIEYKTVFLVFSAVVFGAMALGQTSSFAPDYAKAKISAAHLFVLFNRVPPIDSYR
                     EDGEKPEKFGGNTRIKDVKFNYPNRPEVKILQGLNLAVEKGETLALVGSSGCGKSTVV
                     QLLERFYDPLSGEIVFDDIDAKTLNIQWLRSHIGIVSQEPILFDFTIAENIAYGDNSR
                     EVSHEEIISAAKAASIHSFIDSLPEKYNTRVGDKGTQLSGGQKQRIAIARALIRKPQI
                     LLLDEATSALDTESEKIVQEALDKAREGRTCIVIAHRLSTIQNADKIAVIQNGKVIEQ
                     GTHQQLLAEKGFYYSLVNVQSGSCNM"
     mat_peptide     112..3972
                     /gene="cmdr1"
                     /product="ABC transporter protein"
                     /note="P-glycoprotein"
BASE COUNT     1288 a    758 c    911 g   1061 t
ORIGIN

Alignment Scores:
Pred. No.:              1.56e-39        Length:         4018
Score:                  469.00          Matches:        91
Percent Similarity:     89.43%          Conservative:   19
Best Local Similarity:  73.98%          Mismatches:     13
Query Match:            71.17%          Indels:         0
DB:                     5               Gaps:           0

US-09-873-409-3 (1-131) x GGA9799 (1-4018)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
           :::|||      |||:::::::::|||||||||||:::   |||:::   |||||||||
Db    1489 ATTGACGGGCAGGATCTGAAGAGCCTGAATGTAAGATATCTGCGAGAGATTATTGGTGTG 1548

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           |||:::||||||||||||||||||   ||||||||||:::   ||||||:::|||||||||
Db    1549 GTGAACCAGGAGCCCGTGCTGTTTGCTACAACTATTGCAGAAAATATTCGTTATGGCCGT 1608

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
           :::|||||||||   ||||||:::|||||||||   :::||||||||||||||||||||||
Db    1609 GAGGATGTCACCATGGAAGAAATTGAAAGAGCTACCAAGGAAGCTAATGCTTATGATTTC 1668

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
```

```
             ||||||:::   |||   ||||||   |||:::|||||||||:::||||||||||||||
Db    1669 ATCATGAAGCTCCCCAAGAAGTTTGAAACTGTGGTTGGAGAAAGAGGGGCACAGATGAGT 1728

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           |||||||||||||||||||||||||||||||||||||||||||||   ||||||||||||
Db    1729 GGAGGGCAGAAGCAGCGAATAGCAATTGCCCGTGCTCTGGTTCACAATCCTAAAATTCTT 1788

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
           :::||||||||||||||||||||||||||||:::||||||:::|||   |||||||||||
Db    1789 CTGCTTGATGAGGCAACGTCAGCTTTGGATACTGAGAGTGAATCCGTTGTGCAGGCAGCA 1848

Qy     122 LeuGluLys 124
           |||:::|||
Db    1849 CTGGACAAG 1857

RESULT 15
AF269224
LOCUS       AF269224                4045 bp    DNA     linear   SYN 04-JUL-2000
DEFINITION  Synthetic construct Canis familiaris his-tagged-multidrug
            resistance glycoprotein gene, complete cds.
ACCESSION   AF269224
VERSION     AF269224.1  GI:8926216
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4045)
  AUTHORS   Roulet,A.C., Puel,O., Gesta,S., Alvinerie,M. and Pineau,T.J.
  TITLE     Analysis of dog MDR1 p-glycoprotein
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4045)
  AUTHORS   Roulet,A.C., Puel,O., Gesta,S., Alvinerie,M. and Pineau,T.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-MAY-2000) Pharmacology, INRA, 180 Chemin de
            Tournefeuille, BP3, Toulouse 31931, France
FEATURES             Location/Qualifiers
     source          1..4045
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="original mRNA isolated from Beagle dog colon"
     CDS             101..3964
                     /codon_start=1
                     /transl_table=11
                     /product="his-tagged-multidrug resistance glycoprotein
                     MDR1"
                     /protein_id="AAF81747.1"
                     /db_xref="GI:8926217"
                     /translation="MHHHHHHDPEGGRKGSAEKNFWKMGKKSKKNEKKEKKPTVSTFA
                     MFRYSNWLDRLYMLVGTMAAIIHGAALPLMMLVFGNMTDSFANAGISRNKTFPVIINE
                     SITNNTQHFINHLEEEMTTYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQILKIRKQFF
                     HAIMRQEIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSIATFFTGFIVGFTR
                     GWKLTLVILAISPVLGLSAAIWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFG
                     GQKKELERYNKNLEEAKRIGIKKAITANISIGAAFLLIYASYALAFWYGTSLVLSSEY
                     SIGQVLTVFFSVLIGAFSIGQASPSIEAFANARGAAYEIFKIIDNKPSIDSYSKSGHK
                     PDNIKGNLEFKNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQR
                     LYDPTDGMVCIDGQDIRTINVRHLREITGVVSQEPVLFATTIAENIRYGRENVTMDEI
                     EKAVKEANAYDFIMKLPNKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEAT
                     SALDTESEAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELM
                     KEKGIYFKLVTMQTRGNEIELENATGESKSESDALEMSPKDSGSSLIKRRSTRRSIHA
                     PQGQDRKLGTKEDLNENVPPVSFWRILKLNSTEWPYFVVGIFCAIINGGLQPAFSIIF
                     SRIIGIFTRDEDPETKRQNSNMFSVLFLVLGIISFITFFLQGFTFGKAGEILTKRLRY
                     MVFRSMLRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAVITQNIANLGTGII
                     ISLIYGWQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRT
                     VVSLTREQKFEYMYAQSLQVPYRNSLRKAHIFGVSFSITQAMMYFSYAGCFRFGAYLV
                     ANEFMNFQDVLLVFSAIVFGAMAVGQVSSFAPDYAKAKVSAAHVIMIIEKSPLIDSYS
                     PHGLKPNTLEGNVTFNEVVFNYPTRPDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVV
                     QLLERFYDPLAGSVLIDGKEIKHLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSR
                     VVSHEEIMQAAKEANIHHFIETLPEKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHI
                     LLLDEATSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGKVKEH
                     GTHQQLLAQKGIYFSMISVQAGAKR"
BASE COUNT     1171 a     850 c     986 g    1038 t
ORIGIN

Alignment Scores:
Pred. No.:              1.57e-39      Length:         4045
Score:                  469.00        Matches:        90
Percent Similarity:     89.43%        Conservative:   20
Best Local Similarity:  73.17%        Mismatches:     13
Query Match:            71.17%        Indels:         0
DB:                     12            Gaps:           0

US-09-873-409-3 (1-131) x AF269224 (1-4045)

Qy       2 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 21
           :::|||      |||||||||   :::||||||||||||   |||:::.     ||||||
Db    1481 ATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTG 1540

Qy      22 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 41
           ||||||||||||||||||||||||   |||||||||:::   ||||||:::|||||||||
Db    1541 GTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGC 1600

Qy      42 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 61
           ::::::||||||   :::|||:::|||:::|||   :::|||||||||||||||||||||
Db    1601 GAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTT 1660

Qy      62 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 81
           ||||||:::   ||||||||||||:::||||||||||||||||:::|||||||||:::|||
Db    1661 ATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGT 1720

Qy      82 GlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeu 101
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1721 GGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTT 1780

Qy     102 IleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 121
           :::|||||||||||||||||||||||||||:::||||||::::::   ||||||   |||
Db    1781 CTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCC 1840

Qy     122 LeuGluLys 124
           |||:::|||
Db    1841 CTGGATAAG 1849

Search completed: March 31, 2003, 02:10:06
Job time : 1781.98 secs
```

```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 30, 2003, 03:22:44 ; Search time 7277.11 Seconds
                                           (without alignments)
                                           2354.618 Million cell updates/sec

Title:          US-09-873-409-4
Perfect score:  5330
Sequence:       1 MVISLTSKELSAYSKAGAVA..........QELLRNRDIYFKLVNAQSVQ 1058

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters:      32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115420_19240/app_query.fasta_1.7544
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_30544_@runat_27032003_115420_19240 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      EST:*
                1:  em_estba:*
                2:  em_esthum:*
                3:  em_estin:*
                4:  em_estmu:*
                5:  em_estov:*
                6:  em_estpl:*
                7:  em_estro:*
                8:  em_htc:*
                9:  gb_est1:*
                10:  gb_est2:*
                11:  gb_htc:*
                12:  gb_est3:*
                13:  gb_est4:*
                14:  gb_est5:*
                15:  em_estfun:*
                16:  em_estom:*
                17:  gb_gss:*
                18:  em_gss_hum:*
                19:  em_gss_inv:*
                20:  em_gss_pln:*
                21:  em_gss_vrt:*
                22:  em_gss_fun:*
                23:  em_gss_mam:*
                24:  em_gss_mus:*
                25:  em_gss_other:*
                26:  em_gss_pro:*
                27:  em_gss_rod:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1092	20.5	669	13	BM013981	BM013981 603639619
	2	1051	19.7	1033	14	BM926413	BM926413 AGENCOURT
	3	1008	18.9	943	9	AL520322	AL520322 AL520322
	4	995	18.7	760	13	BM016204	BM016204 603642659
	5	918	17.2	545	9	AL040762	AL040762 DKFZp434C
	6	903	16.9	929	14	BQ882401	BQ882401 AGENCOURT
	7	856.5	16.1	948	13	BM471690	BM471690 AGENCOURT
	8	834	15.6	998	14	BM904842	BM904842 AGENCOURT
	9	780	14.6	2676	11	AK014319	AK014319 Mus muscu
	10	771	14.5	894	12	BF584668	BF584668 602098406
	11	770.5	14.5	1019	12	BG248052	BG248052 602359987
	12	753	14.1	726	12	BG293345	BG293345 602390738
	13	752.5	14.1	871	17	AZ682350	AZ682350 ENTKB16TF
	14	751	14.1	944	12	BF796582	BF796582 602258463
	15	743.5	13.9	886	17	BH139685	BH139685 ENTNA47TR
c	16	734.5	13.8	932	17	AZ670821	AZ670821 ENTJN69TF
	17	728.5	13.7	947	17	AZ683753	AZ683753 ENTIL96TF
	18	726.5	13.6	886	17	AZ540627	AZ540627 ENTEQ18TF
c	19	723.5	13.6	880	17	AZ687805	AZ687805 ENTLU52TF
	20	721.5	13.5	939	14	BQ720763	BQ720763 AGENCOURT
	21	720.5	13.5	897	17	AZ541090	AZ541090 ENTDS67TR
	22	719	13.5	780	12	BG587938	BG587938 EST489713
	23	711.5	13.3	913	17	BH155700	BH155700 ENTRO54TR
	24	709.5	13.3	834	17	AZ548312	AZ548312 ENTFG07TR
c	25	708.5	13.3	853	17	AZ679807	AZ679807 ENTHI67TR
	26	705.5	13.2	823	17	AZ532602	AZ532602 ENTCR50TR
	27	697.5	13.1	861	14	BQ717101	BQ717101 AGENCOURT
	28	692.5	13.0	1341	11	AY108285	AY108285 Zea mays
	29	690.5	13.0	899	17	BH154857	BH154857 ENTRO47TF
c	30	690	12.9	835	13	BJ349604	BJ349604 BJ349604
c	31	689.5	12.9	891	17	AZ682250	AZ682250 ENTKT68TR
	32	688	12.9	1123	11	AY108485	AY108485 Zea mays
	33	687	12.9	715	14	BQ869512	BQ869512 QGD6G09.y
c	34	684	12.8	820	13	BJ356540	BJ356540 BJ356540
	35	683	12.8	600	13	BM486593	BM486593 pgm2n.pk0
	36	682.5	12.8	919	17	AZ690701	AZ690701 ENTIY68TR
	37	680	12.8	926	14	BQ123477	BQ123477 EST609053
	38	677.5	12.7	795	12	BF313560	BF313560 601900192
	39	674.5	12.7	815	13	BI409057	BI409057 602961127
	40	670	12.6	852	17	BH720383	BH720383 BOHZZ94TF
	41	669	12.6	547	13	BJ030711	BJ030711 BJ030711
c	42	667	12.5	937	17	BH150760	BH150760 ENTQG42TR
c	43	666	12.5	792	13	BJ352404	BJ352404 BJ352404
c	44	663	12.4	838	13	BJ355307	BJ355307 BJ355307
	45	662	12.4	636	10	AV962688	AV962688 AV962688

ALIGNMENTS

```
RESULT 1
BM013981
LOCUS       BM013981                 669 bp    mRNA    linear   EST 30-OCT-2001
DEFINITION  603639619F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5415556 5',
            mRNA sequence.
ACCESSION   BM013981
VERSION     BM013981.1  GI:16528335
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 669)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
```

```
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12060  row: f  column: 05
            High quality sequence stop: 664.
FEATURES             Location/Qualifiers
     source          1. .669
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5415556"
                     /clone_lib="NIH_MGC_87"
                     /tissue_type="mammary adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.383 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      203 a    122 c    162 g    182 t
ORIGIN

Alignment Scores:
Pred. No.:               8.28e-115     Length:        669
Score:                   1092.00       Matches:       215
Percent Similarity:      99.54%        Conservative:  2
Best Local Similarity:   98.62%        Mismatches:    0
Query Match:             20.49%        Indels:        1
DB:                      13            Gaps:          0

US-09-873-409-4 (1-1058) x BM013981 (1-669)

Qy    128 AlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPhe 147
          :::||||||||:::||||||||||||||||||||||||||||||||||||||||||||||
Db     17 AGTGTTTTCTTTAATGTAATCCATAGCAGTTATTGCATTGGAGCAGCAGTCCCTCACTTT 76

Qy    148 GluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLys 167
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     77 GAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTTATTGATAAGAAA 136

Qy    168 ProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrVal 187
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    137 CCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTG 196

Qy    188 GluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGly 207
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    197 GAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGT 256

Qy    208 LeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGly 227
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    257 CTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGG 316

Qy    228 LysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMet 247
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    317 AAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATG 376

Qy    248 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 267
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    377 GTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTG 436

Qy    268 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 287
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    437 GTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGA 496

Qy    288 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 307
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    497 GATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTT 556

Qy    308 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 327
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    557 ATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATGAGT 616

Qy    328 GlyGlyGlnLysGlnArg-IleAlaIleAlaArgAlaLeuValArgAsnPro 344
          |||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    617 GGAGGGCAGAAACAGAGGCATCGCAATTGCTCGTGCCTTAGTTCGAAACCCC 668

RESULT 2
BM926413
LOCUS       BM926413                1033 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6600788 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5764845
            5', mRNA sequence.
ACCESSION   BM926413
VERSION     BM926413.1  GI:19376792
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1033)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12818  row: o  column: 22
            High quality sequence start: 3
            High quality sequence stop: 686.
FEATURES             Location/Qualifiers
     source          1. .1033
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5764845"
                     /clone_lib="NIH_MGC_114"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
                     male brains, age range 23-27 yo. Library is oligo-dT
                     primed and directionally cloned (EcoRV site is  destroyed
                     upon cloning).  Average insert size 1.5 kb, insert size
                     range 1-3 kb.  Library is normalized and enriched for
                     full-length  clones and was constructed by C. Gruber
                     (Invitrogen).  Research  Genetics tracking code 019. Note:
                     this is a NIH_MGC Library."
BASE COUNT      289 a    212 c    218 g    312 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:               9.2e-110      Length:        1033
Score:                   1051.00       Matches:       221
Percent Similarity:      94.49%        Conservative:  2
Best Local Similarity:   93.64%        Mismatches:    13
Query Match:             19.72%        Indels:        3
DB:                      14            Gaps:          0

US-09-873-409-4 (1-1058) x BM926413 (1-1033)

Qy    129 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 148
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 GTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAA 239

Qy    149 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 168
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 ACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTCATTGATAAGAAACCC 299
```

```
Qy     169 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 188
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     300 AGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGAA 359

Qy     189 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 208
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     360 TTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCTG 419

Qy     209 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 228
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     420 AATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAG 479

Qy     229 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 248
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     480 AGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGTG 539

Qy     249 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 268
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     540 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTGGTT 599

Qy     269 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 288
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     600 AGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGAGAT 659

Qy     289 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 308
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     660 GATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATC 719

Qy     309 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 328
           |||||||||||||||    ||||||||||||||||||||     ||||||   ||||||||
Db     720 ATGGAGTTTCCTAATAA-TTTAATACATTGGTAGGGGAANAAGGAGCTCCAATGAGTGGA 778

Qy     329 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 348
              |||     ||||||   ||||||   ||||||||||||||||||:::   |||
Db     779 AGCCAGAACCG-AGGATCCCAATTGCTCCTGCCTTAGTTCGAAACCCCAGGAT-CTGAAT 836

Qy     349 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 364
           ||||||||||||||||||||||||||||||||:::        |||
Db     837 TTAGATGAAGCTACGTCTGCCCTGGATTCAGAAACCAGTCAACTGGTC 884

RESULT 3
AL520322
LOCUS       AL520322                943 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL520322 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB006YC15 5
            prime, mRNA sequence.
ACCESSION   AL520322
VERSION     AL520322.1  GI:12783815
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 943)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1..943
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DB006YC15"
                     /clone_lib="LTI_NFL004_NBC2"
                     /sex="male"
                     /tissue_type="neuroblastoma cells"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed
                     by Life Technologies. Contact : Feng Liang Life
                     Technologies, a division of Invitrogen 9800 Medical Center
                     Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
                     8371 Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      253 a    225 c    253 g    211 t    1 others
ORIGIN

Alignment Scores:
Pred. No.:              7.11e-105      Length:         943
Score:                  1008.00        Matches:        194
Percent Similarity:     80.40%         Conservative:   48
Best Local Similarity:  64.45%         Mismatches:     58
Query Match:            18.91%         Indels:         1
DB:                     9              Gaps:           0

US-09-873-409-4 (1-1058) x AL520322 (1-943)

Qy     757 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 776
           :::::|||:::|||:::   :::|||||||||||:::|||:::              |||
Db       5 TTATTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCT 64

Qy     777 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 796
           |||:::|||:::|||||||||      |||||||||:::         ::::::||||||
Db      65 CCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACC 124

Qy     797 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 816
           |||   |||||||||   |||   ||||||      |||:::|||   |||||||||:::
Db     125 CCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTC 184

Qy     817 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 836
              |||   ||||||   |||   ||||||   |||||||||:::   :::|||:::|||
Db     185 ACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGA 244

Qy     837 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 856
           |||||||||   :::::::::|||:::|||:::|||   ||||||||||||||||||||||
Db     245 CTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGG 304

Qy     857 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 876
           |||||||||   ||||||||||||:::|||   |||||||||:::   |||:::||||||
Db     305 AAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTG 364

Qy     877 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 896
              ||||||   :::   |||   ||||||||||||||||||||||:::   :::   |||
Db     365 CTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATC 424

Qy     897 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 916
           |||   |||||||||:::||||||:::|||||||||||||||||||||||||||||||||
Db     425 GTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGAC 484

Qy     917 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 936
           |||||||||||||||         ||||||      ||||||   |||||||||||||
Db     485 AACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACAT 544

Qy     937 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 956
           :::|||||||||   ||||||   ||||||:::|||:::||||||   ||||||   |||
Db     545 GCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAG 604

Qy     957 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 976
           ||||||||||||||||||||||||:::||||||||||||||||||:::::::::|||
Db     605 CTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCAT 664

Qy     977 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 996
           |||||||||||||||||||||||||||||||||   :::|||||||||||||||||||
Db     665 ATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAA 724

Qy     997 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 1016
              |||||||||||||||||   ||||||||||||:::|||:::   ||||||||||||
Db     725 GAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCC 784

Qy    1017 AlaIleGlnAsnAlaAspLeuIleVal-ValLeuHisAsnGlyLysIleLysGluGlnGl 1036
```

```
             |||||||||||||||||||||| |||       ||||||::::::||||||    ||
Db     785 ACCATCCAGAATGCAGACTTAATAGTGTGTGTTTCAGAATGGCAGAGTCAAGGAGCATGG 844

Qy    1036 yThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSe 1056
             |||||||||:::||||||        :::   |||||||||   :::|||:::    |||::
Db     845 CACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTCAATGGTCAGTGTCCAGGC 904

Qy    1056 r 1056
           :
Db     905 T 905

RESULT 4
BM016204
LOCUS       BM016204                 760 bp    mRNA    linear   EST 30-OCT-2001
DEFINITION  603642659F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5418615 5',
            mRNA sequence.
ACCESSION   BM016204
VERSION     BM016204.1  GI:16530558
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 760)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12068  row: e  column: 16
            High quality sequence stop: 740.
FEATURES             Location/Qualifiers
     source          1..760
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5418615"
                     /clone_lib="NIH_MGC_87"
                     /tissue_type="mammary adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.383 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      229 a    150 c    187 g    193 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              1.55e-103       Length:         760
Score:                  995.00          Matches:        203
Percent Similarity:     98.07%          Conservative:   0
Best Local Similarity:  98.07%          Mismatches:     2
Query Match:            18.67%          Indels:         2
DB:                     13              Gaps:           0

US-09-873-409-4 (1-1058) x BM016204 (1-760)

Qy     166 LysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGly 185
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3 AAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGA 62

Qy     186 ThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeu 205
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63 ACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTG 122

Qy     206 LysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGly 225
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 AAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGC 182

Qy     226 SerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPhe 245
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 AGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTT 242

Qy     246 IleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle 265
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 ATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATN 302

Qy     266 GlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr 285
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 GGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTAT 362

Qy     286 GlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyr 305
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     363 GGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTAT 422

Qy     306 AspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGln 325
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     423 GATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAA 482

Qy     326 MetSerGlyGlyGlnLysGlnArgIleAlaIle-AlaArgAlaLeuValArgAsnProLy 345
             |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db     483 ATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGGCTCGTGCCTTAGTTCGAAACCCCAA 542

Qy     345 sIleLeuIleLeu-AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValG 365
             |||||||||    ||||||||||||||||||||||||||||||||||||||||||||||
Db     543 GATTCTGATTTTCAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTC 602

Qy     365 lnAlaAlaLeuGluLys 370
             ||||||||||||||||
Db     603 AAGCTGCACTGGAGAAG 619

RESULT 5
AL040762
LOCUS       AL040762                 545 bp    mRNA    linear   EST 29-FEB-2000
DEFINITION  DKFZp434C1815_r1 434 (synonym: htes3) Homo sapiens cDNA clone
            DKFZp434C1815 5', mRNA sequence.
ACCESSION   AL040762
VERSION     AL040762.1  GI:5409708
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 545)
  AUTHORS   Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
            Wiemann,S.
  TITLE     EST (Bloecker, et al.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Bloecker H
            MIPS
            Am Klopferspitz 18a D-82152 Martinsried, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by GBF (National Research Centre for Biotechnology Ltd.,
            Braunschweig/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            s1 sequence also available.
            This clone (DKFZp434C1815) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1..545
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp434C1815"
                     /clone_lib="434 (synonym: htes3)"
                     /tissue_type="testis"
                     /dev_stage="adult"
```

```
                  /lab_host="DH10B"
                  /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT      171 a     89 c    138 g     147 t
ORIGIN

Alignment Scores:
Pred. No.:               6.58e-95      Length:        545
Score:                   918.00        Matches:       181
Percent Similarity:      100.00%       Conservative:  0
Best Local Similarity:   100.00%       Mismatches:    0
Query Match:             17.22%        Indels:        0
DB:                      9             Gaps:          0

US-09-873-409-4 (1-1058) x AL040762 (1-545)

Qy    169 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 188
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3 AGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGAA 62

Qy    189 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 208
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     63 TTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCTG 122

Qy    209 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 228
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    123 AATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAG 182

Qy    229 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 248
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    183 AGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGTG 242

Qy    249 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 268
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    243 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTGGTT 302

Qy    269 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 288
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    303 AGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGAGAT 362

Qy    289 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 308
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    363 GATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATC 422

Qy    309 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 328
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    423 ATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGA 482

Qy    329 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 348
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    483 GGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATT 542

Qy    349 Leu 349
          |||
Db    543 TTA 545

RESULT 6
BQ882401
LOCUS       BQ882401                 929 bp    mRNA    linear   EST 16-AUG-2002
DEFINITION  AGENCOURT_8627902 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6291782
            5', mRNA sequence.
ACCESSION   BQ882401
VERSION     BQ882401.1  GI:22274409
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 929)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2492  row: c  column: 15
            High quality sequence stop: 677.
FEATURES             Location/Qualifiers
     source          1..929
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6291782"
                     /clone_lib="NIH_MGC_43"
                     /tissue_type="normal pigmented retinal epithelium"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Library constructed  by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and  Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library. |"
BASE COUNT      281 a    176 c    237 g    234 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:               8.72e-93      Length:        929
Score:                   903.00        Matches:       188
Percent Similarity:      73.02%        Conservative:  61
Best Local Similarity:   55.13%        Mismatches:    58
Query Match:             16.94%        Indels:        34
DB:                      14            Gaps:          3

US-09-873-409-4 (1-1058) x BQ882401 (1-929)

Qy     12 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThr-ValIl 31
          ||||||:::|||||||||||||||||||||||||||||::::::|||||||| |||||
Db      1 GCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGGTGAT 60

Qy     31 eAlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTy 51
           ||||||    |||:::||||||||:::
Db     61 TGCATTTGGAGGACAAAAGAAAGAACTTGAA---------------------------- 91

Qy     51 rAlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrTh 71
                                                             ||||||
Db     92 --------------------------------------------------AGGTACAA 99

Qy     71 rGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSe 91
           :::||||||::::::||||||       ||||||||:::   |||    :::   :::||
Db    100 CAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTC 159

Qy     91 rLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyTh 111
          |:::||||||   :::       :::      :::|||  ||||||||||||||||||
Db    160 TATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGAC 219

Qy    111 rSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePh 131
          |:::|||:::|||:::||||||       ||||:::||||||   |||||||  ||||||||
Db    220 CACCTTGGTCCTCTCAGGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTT 273

Qy    131 eSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAl 151
          |||||||:::        ::::::   |||   |||         |||    ||||||
Db    274 TTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGC 333

Qy    151 aIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAs 171
          |   ||||||||||||||:::    ||||||::::::||||||    |||||||||||||
Db    334 AAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGA 393

Qy    171 pAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAs 191
          |::::::|||    :::|||:::||||||::::::|||:::|||    :::||||||:::||
Db    394 CAGCTATTCGAAGAGTGGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAA 453
```

```
Qy    191 nValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuAr 211
Db    454 TGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAGATCTTGAAGGGTCTGAACCTGAA 513

Qy    211 gIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVa 231
Db    514 GGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAAC 573

Qy    231 lValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAs 251
Db    574 AGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACA 633

Qy    251 nAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGl 271
Db    634 GGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGA 693

Qy    271 uProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValTh 291
Db    694 ACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCCCTATGGCCGTGAAAATGTCAC 753

Qy    291 rAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla-TyrAspPheIleMetGluP 311
Db    754 CATGGATGAGATTGAGAAAGCTGTCCAGGAAGCCAATGCCCTATGACTTTATCATGAAAC 813

Qy    311 heProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGly---G 330
Db    814 TGCCTCATAAATTTGACACCCTGGTTGGANAAAGAAGGGGCCAATTGAGTGGTGGGGCAC 873

Qy    330 lnLysGlnArgIleAlaIleAlaArg-AlaLeuVal-ArgAsnProLysIleLeu 347
Db    874 AAACCAAAGGATCTCCCTTGGCACGTGGCCCTGGTTCCCCAACCCCAAGATCCTC 928
```

```
RESULT 7
BM471690
LOCUS       BM471690                 948 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6465349 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5539117
            5', mRNA sequence.
ACCESSION   BM471690
VERSION     BM471690.1  GI:18520732
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 948)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12233  row: b  column: 14
            High quality sequence stop: 569.
FEATURES             Location/Qualifiers
     source          1..948
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5539117"
                     /clone_lib="NIH_MGC_72"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.  Library constructed by Life
                     Technologies."
BASE COUNT      285 a    187 c    194 g    282 t
ORIGIN

Alignment Scores:
Pred. No.:                2.06e-87       Length:          948
Score:                    856.50         Matches:         186
Percent Similarity:       89.67%         Conservative:    5
Best Local Similarity:    87.32%         Mismatches:      14
Query Match:              16.07%         Indels:          8
DB:                       13             Gaps:            1

US-09-873-409-4 (1-1058) x BM471690 (1-948)
```

```
Qy    129 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 148
Db    205 GTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAA 264

Qy    149 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 168
Db    265 ACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTTATTGATAAGAAACCC 324

Qy    169 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 188
Db    325 AGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGAA 384

Qy    189 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 208
Db    385 TTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCTG 444

Qy    209 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 228
Db    445 AATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAG 504

Qy    229 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 248
Db    505 AGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGTG 564

Qy    249 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle-GlyValVa 268
Db    565 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTTGGAGTGGT 624

Qy    268 lSerGlnGluPro--ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr-GlyAr 287
Db    625 TAGTCAAGAGCCCTGGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGGACG 684

Qy    287 gAspAspValThrAspGluGlu-MetGluArgAlaAlaArgGluAla-AsnAlaTyrAsp 306
Db    685 AGATGATGGGACTGATGAACAAAATGGAAAAAGCAACAAGGGAAACCAAAGGCGAATGAA 744

Qy    307 PheIleMetGlu-PheProAsnLysPheAsnThrLeuVal---GlyGluLysGlyAlaGl 325
Db    745 TTTATCATGGAATTTCCCAAAAAAATTTAATACATTGGAAAGGGGAAAAAAGGACCTTCA 804

Qy    325 nMetSerGlyGlyGlnLysGlnArg 333
Db    805 AGGAATGGAAGGGCCAAAACCCAAG 829
```

```
RESULT 8
BM904842
LOCUS       BM904842                 998 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6699581 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5557655
            5', mRNA sequence.
ACCESSION   BM904842
VERSION     BM904842.1  GI:19355221
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 998)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
```

```
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: ATCC/DCTD/DTP
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Agencourt Bioscience Corporation
           Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM12279  row: f  column: 24
          High quality sequence stop: 738.
FEATURES             Location/Qualifiers
     source          1. .998
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5557655"
                     /clone_lib="NIH_MGC_72"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      271 a    223 c    248 g    256 t
ORIGIN

Alignment Scores:
Pred. No.:              8.81e-85      Length:          998
Score:                  834.00        Matches:         164
Percent Similarity:     74.12%        Conservative:    68
Best Local Similarity:  52.40%        Mismatches:      77
Query Match:            15.65%        Indels:          4
DB:                     14            Gaps:            0

US-09-873-409-4 (1-1058) x BM904842 (1-998)

Qy     581 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600
               |||::::::||||||||:::    :::|||:::||||||   ||||||||   ||||||
Db      61 GATGTGAGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCC 120

Qy     601 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620
                  |||    ||||||::::::||||||    |||||||||:::    |||:::|||||||||
Db     121 AATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAAT 180

Qy     621 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 640
                       |||:::|||        :::|||||||||||||||||||||||||||::::::|||
Db     181 ATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTG 240

Qy     641 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 660
               |||:::|||:::|||    |||::::::|||:::    |||::::::|||             :::
Db     241 TTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTG 300

Qy     661 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 680
               :::|||    |||    ||||||||||:::||||||:::    ||||||||||||||||||||
Db     301 TCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAA 360

Qy     681 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 700
               |||:::||||||    ||||||:::||||||||||||:::|||:::    ||||||    |||
Db     361 GCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATG 420

Qy     701 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720
               |||    :::    ||||||          :::||||||:::    :::||||||    |||    |||
Db     421 TATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGA 480

Qy     721 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 740
                  ::::::|||:::    |||    :::|||||||:::||||||         |||||||||
Db     481 ATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTT 540

Qy     741 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760
               ||||||||||||:::             |||:::    |||    :::    :::||||||:::
Db     541 GGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCA 600

Qy     761 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 780
               |||:::    :::||||||||||||:::|||:::               ||||||:::|||:::
Db     601 GCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCC 660

Qy     781 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 800
               ||||||||        |||||||||:::         ::::::||||||    |||    ||||||
Db     661 AAAGCCAAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGAC 720

Qy     801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820
               |||    |||    ||||||         |||:::|||    |||||||||:::    |||
Db     721 AGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGGTGA 780

Qy     821 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer-LeuSe 840
               |||    |||    ||||||    ||||||||:::    :::|||:::||||||||| |||
Db     781 GTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCCTGGA 840

Qy     840 rIleGluArgGlyLysThrValAlaPheVal-GlySerSerGlyCysGlyLysSerThrS 860
             :::::::::|||:::|||:::|||    ||| ||||||||||||||||||||||||
Db     841 GGTGAAGAAGGGGCAGACCCTGGCTCTGGTGGGGCAGCAGTGGCTGTGGGAAGAGCACAG 900

Qy     860 er-ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 879
                  ||||||||||:::|||:::|||:::|||:::         ::::::|||    ||||||
Db     901 TGGGTCCAGCTCCTGGAGCGGGTCTACAACCCCTTGGCAAGGAAAATGCTGCTTGATGGC 960

Qy     880 ValAspAlaLys-GluLeuAsnValGlnTrp 889
                         |||    |||:::|||    |||
Db     961 AAAAGAATAAAACCGACTGGATGTTCCATGG 991

RESULT 9
AK014319
LOCUS       AK014319              2676 bp    mRNA    linear   HTC 19-JAN-2002
DEFINITION  Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length
            enriched library, clone:3222401P09:ATP-binding cassette, sub-family
            B (MDR/TAP), member 8, full insert sequence.
ACCESSION   AK014319
VERSION     AK014319.1  GI:12852089
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (strain:C57BL/6J) 14, 17 days embryo head cDNA to
            mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:3222401P09.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
```

```
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409 (6821), 685-690 (2001)
  MEDLINE   21085660
   PUBMED   11217851
REFERENCE   5  (bases 1 to 2676)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C.,
            Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,
            Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F.,
            Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,
            Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,
            Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
            Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,
            Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and
            Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Gemomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. Second
            strand cDNA was prepared with the primer adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cleaved
            with BamHI and XhoI. cDNA of size comprised between 3 and 7 kb was
            selected before cloning. Vector: a modified pBluescript KS(+) after
            bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'
            end: BamHI. Host: DH10B.
FEATURES             Location/Qualifiers
     source          1..2676
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:3222401P09"
                     /db_xref="MGD:MGI:1907488"
                     /db_xref="taxon:10090"
                     /clone="3222401P09"
                     /tissue_type="head"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="14, 17 days embryo"
     gene            1..2676
                     /gene="Abcb8"
     CDS             72..2225
                     /gene="Abcb8"
                     /note="ATP-binding cassette, sub-family B (MDR/TAP),
                     member 8
                     data source:MGD, source key:MGI:1351667, evidence:ISS
                     putative"
                     /codon_start=1
                     /protein_id="BAB29270.1"
                     /db_xref="GI:12852090"
                     /db_xref="MGD:MGI:1351667"
                     /translation="MLVHLFRFGIRGGPVPGWSLQSLRFQTFSAARSSDDRLSSHLLR
                     TVAQLRVQLRAHLPRAPPASHWSPSAWCWVGGTLVVPAVLWQHPRLCLIALCEAKESP
                     PAQPTRAPELRFNWKLFWHFLHPHLLALGAAIVLALGAALVNVQIPLLLGQLVEIVAK
                     YTRDHMGSFVSESRKLSVQLLLLYGVQGLLTFGYLVLLSHIGERMAMDMRKALFSSLL
                     RQDIAFFDAKKTGQLVSRLTTDVQEFKSSFKLVISQGLRSCTQVIGSLVSLSMLSPRL
                     TLMLAVVTPALMGVGTLMGSGLRKLSRQCQEQIARATGVADEALGNVRTVRAFAMEKR
                     EEERYQAELESCCCKAEELGRGIALFQGLSNIAFNCMVLGTLFIGGSLVAGQQLKGGD
                     LMSFLVASQTVQRSMASLSVLFGQVVRGLSAGARVFEYMALSPVIPLTGGYCIPNKDI
                     RGSITFQNVTFSYPCRPGFNVLKDFTLKLPSGKIVALVGQSGGGKTTVASLLERFYDP
                     EAGSVTLDGHDLRTLNPSWLRGQVIGFISQEPVLFATTIMENIRFGKLDASDEEVYTA
                     AREANAHEFISSFPDGYSTVVGERGTTLSGGQKQRLAIARALIKQPTVLILDEATSAL
                     DAESERVVQEALDRASAGRTVLVIAHRLSTVRAAHSIIVMANGQVCEAGTHEELLKKG
                     GLYSELIRRQTLDASLTSTPPAEKPEDPKSCQSKA"
     polyA_signal    2653..2658
                     /gene="Abcb8"
                     /note="putative"
     polyA_site      2676
                     /gene="Abcb8"
                     /note="putative"
BASE COUNT      544 a    785 c    756 g    591 t
ORIGIN

Alignment Scores:
Pred. No.:              7.87e-78      Length:         2676
Score:                  780.00        Matches:        207
Percent Similarity:     51.28%        Conservative:   113
Best Local Similarity:  33.17%        Mismatches:     250
Query Match:            14.63%        Indels:         54
DB:                     11            Gaps:           14

US-09-873-409-4 (1-1058) x AK014319 (1-2676)

Qy        462 LysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLys 481
              :::|||:::||||||      ::::::            ||||||
Db        360 GAGGCCAAAGAGTCCCCTCCTGCCCAGCCCACACGTGCCCCAGAG--------------- 404

Qy        482 IleLeuLysLeuAsn---LysProGluTrpProPhe------------------------ 492
                 |||:::   |||    |||      |||   |||
Db        405 ---CTGCGCTTTAATTGGAAGCTCTTCTGGCATTTCCTTCACCCCCACCTGCTGGCCTTG 461

Qy        493 ---------ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProVal 509
                       :::||||||      |||:::::::::|||        :::    |||:::
Db        462 GGGGCAGCCATTGTGCTGGCCTTAGGTGCGGCACTAGTGAATGTGCAGATC---CCCTTG 518

Qy        510 PheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeu 529
                                 :::|||:::    :::      :::          :::
Db        519 CTCCTGGGCCAGCTGGTGGAGATTGTCGCCAAGTACACGAGGGACCACATGGGGAGTTTC 578

Qy        530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
                 ::::::             |||:::      ::::::|||
Db        579 GTGTCTGAGTCCCGTAAGCTCAGCGTCCAGCTGCTCCTACTG------------------ 620

Qy        550 SerTyrPheMetGlnGlyLeu---------------------PheTyrGlyArgAlaGly 562
                 |||    :::|||||||||                                   |||
Db        621 ---TACGGTGTTCAGGGACTGCTGACCTTTGGATACCTAGTGCTGCTGTCCCACATTGGT 677

Qy        563 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 582
              |||   :::   |||   :::|||         |||   ::::::|||   ||||||||
Db        678 GAGCGCATGGCCATGGACATGCGGAAAGCCCTTTTCAGCTCCCTGCTCCGGCAAGACATT 737

Qy        583 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 602
              |||:::||||||   |||:::        ||||||   |||   :::   |||    |||
Db        738 GCTTTCTTTGATGCCAAAAAG------ACAGGGCAGCTAGTGAGTCGCTTGACTACTGAT 791

Qy        603 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 622
```

```
               :::   :::   :::   :::              |||            :::::  |||
Db    792 GTGCAAGAATTCAAGTCATCCTTCAAGCTTGTCATCTCCCAGGGACTGCGCAGCTGCACC 851

Qy    623 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 642
              :::      |||::::::        :::          :::|||   ::::::
Db    852 CAGGTGATTGGTAGCCTGGTGTCCCTGTCTATGCTGTCCCCTCGCCTTACCCTGATGCTG 911

Qy    643 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 662
              :::   |||  |||         |||   :::        :::  :::
Db    912 GCTGTCGTCACACCCGCCCTCATGGAGTGGGCACCCTGATGGCTCAGGCCTCCGAAAG 971

Qy    663 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 682
              :::   :::   ::::::::::::        |||      :::|||  |||||||||
Db    972 CTGTCTCGCCAGTGTCAGGAGCAGATTGCCAGGGCAACAGGTGTAGCAGATGAGGCCCTT 1031

Qy    683 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 702
             |||:::||||||:::   :::         |||||      |||:::  |||:::
Db   1032 GGCAATGTTCGGACTGTGCGGGCCTTCGCGATGGAGAAGAGGGAGGAGGAACGCTATCAA 1091

Qy    703 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 722
             |||:::                :::      ||||||:::   :::|||
Db   1092 GCAGAACTGGAG------------TCATGCTGCTGTAAAGCAGAAGAACTGGGCAGGGGC 1139

Qy    723 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyr---------AlaAlaGlyPheArg 739
             |||                         |||:::                |||
Db   1140 ATCGCCTTGTTCCAAGGGCTCTCCAACATCGCTTTCAACTGTATGGTCTTGGGCACCCTG 1199

Qy    740 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGly---------MetPhe 756
          |||            :::  ||||||:::                |||         |||
Db   1200 TTCATTGGGGGCTCCCTTGTGGCTGGACAGCAGCTGAAAGGGGGAGACCTCATGTCCTTC 1259

Qy    757 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 776
          :::|||    :::                :::||||||             ||||||
Db   1260 CTGGTGGCTTCCCAGACAGTACAGAGGTCTATGGCC------AGCCTCTCTGTCCTGTTT 1313

Qy    777 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 796
             :::      :::      |||      |||   :::|||       :::
Db   1314 GGTCAGGTGGTACGTGGGCTGAGTGCGGGAGCCCGAGTCTTCGAATACATGGCCCTGAGC 1373

Qy    797 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 816
          |||   |||                                      |||::::::
Db   1374 CCTGTCATCCCATTGACCGGGGGCTACTGCATCCCCAACAAGGACATTCGTGGTTCCATC 1433

Qy    817 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 836
             |||:::   |||:::|||   ||||||||||||||          :::|||:::
Db   1434 ACCTTCCAAAATGTCACCTTCAGCTACCCCTGCAGACCTGGCTTCAATGTGCTCAAGGAC 1493

Qy    837 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 856
             :::|||   :::      ||||||   ||||||   ||||||   ||||||   |||
Db   1494 TTCACCCTGAAGCTGCCCTCTGGCAAGATTGTGGCTCTTGTGGGCCAGTCTGGGGGAGGA 1553

Qy    857 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 876
          |||:::|||         ||||||:::|||   |||||||||      |||   |||
Db   1554 AAGACCACAGTTGCCTCCTTGCTGGAACGCTTCTATGACCCTGAAGCTGGCTCGGTGACG 1613

Qy    877 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGln---IleAla 895
          ||||||   |||   :::   ||||||      |||||||||   |||   |||
Db   1614 TTGGATGGGCATGACCTGCGAACTCTCAACCCCTCCTGGCTCCGGGGCCAGGTCATAGGT 1673

Qy    896 IleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGly 915
             :::   ||||||||||||||||||      :::|||   |||||||||   :::|||
Db   1674 TTCATCAGCCAGGAGCCAGTCCTGTTTGCAACAACCATCATGGAGAATATCCGATTTGGG 1733

Qy    916 ------AspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAla 933
                   |||   |||                :::|||:::      ||||||   |||
Db   1734 AAGCTGGATGCTTCCGAT------------GAAGAGGTGTACACAGCTGCACGAGAAGCC 1781

Qy    934 AsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLys 953
          |||   |||   ||||||         |||:::   |||:::|||   ||||||   :::
Db   1782 AATGCCCACGAGTTCATCAGCAGCTTCCCCGATGGCTACAGCACTGTGGTTGGTGAGCGG 1841

Qy    954 GlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGln 973
          |||   ||||||||||||||||||||||||||||||||||||||||||||||::::::
Db   1842 GGCACAACCTTGTCTGGTGGCCAGAAGCAGCGCCTAGCCATCGCACGTGCCCTCATCAAG 1901

Qy    974 LysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLys 993
          :::|||   :::|||:::||||||||||||||||||||||||      :::||||||:::
Db   1902 CAGCCCACAGTGCTGATCCTGGACGAGGCCACCAGTGCGCTAGATGCAGAATCCGAGAGG 1961

Qy    994 ValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHis 1013
          |||||||||   |||||||||:::|||      |||||||||   ||||||:::   |||
Db   1962 GTGGTACAGGAGGCCCTGGACCGGGCCAGTGCTGGCCGCACCGTGTTGGTCATTGCCCAC 2021

Qy   1014 ArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLys 1033
          |||||||||   ::::::   |||      |||:::|||:::   ||||||::::::
Db   2022 CGGCTTAGTACTGTCCGTGCAGCCCACTCCATCATTGTCATGGCCAATGGCCAAGTCTGT 2081

Qy   1034 GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsn 1053
          |||   |||||||||:::|||||||||:::            :::|||   :::|||:::
Db   2082 GAGGCTGGGACCCACGAAGAACTCCTTAAAAAAGGCGGGCTCTATTCAGAGCTTATCCGG 2141

Qy   1054 AlaGlnSerVal 1057
             |||::::::
Db   2142 AGACAAACCCTG 2153
```

RESULT 10
BF584668

```
LOCUS       BF584668                 894 bp    mRNA    linear   EST 12-DEC-2000
DEFINITION  602098406F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4218385 5',
            mRNA sequence.
ACCESSION   BF584668
VERSION     BF584668.1  GI:11658386
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 894)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9798  row: 1  column: 02
            High quality sequence stop: 651.
FEATURES             Location/Qualifiers
     source          1..894
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4218385"
                     /clone_lib="NCI_CGAP_Co24"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.6 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      263 a    203 c    242 g    185 t     1 others
ORIGIN

Alignment Scores:
Pred. No.:              1.33e-77       Length:         894
Score:                  771.00         Matches:        171
Percent Similarity:     79.72%         Conservative:   57
Best Local Similarity:  59.79%         Mismatches:     48
Query Match:            14.47%         Indels:         15
DB:                     12             Gaps:           3

US-09-873-409-4 (1-1058) x BF584668 (1-894)
```

```
Qy     129 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 148
           ||||||||||||||:::       ::::::   :::|||   |||   |||:::   |||
Db       5 GTCTTCTTTTCCGTGTTAAT-GGAGCATTCAGTGTTGGACAGGCATCTCCAAATATTGAA 63

Qy     149 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 168
              ||||||   |||||||||||||||:::   :::|||::::::||||||   ||||||
Db      64 GCCTTCGCCAATGCACGAGGAGCAGCTTATGAAGTCTTCAAAATAATTGATAATAAGCCC 123

Qy     169 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 188
           |||||||||:::||||||   :::|||:::||||||::::::|||:::|||   :::|||
Db     124 AGTATAGACAGCTTCTCAAAGAGTGGGCACAAACCAGACAACATACAAGGAAATCTGGAA 183

Qy     189 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 208
           |||||||||:::   |||:::||||||||||||      ::::::||||||||||||||||
Db     184 TTTAAGAATATTCACTTCAGTTACCCATCTCGAAAAGAAGTTCAGATCTTGAAGGGCCTC 243

Qy     209 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 228
           ||||||:::::|||||||||:::||||||||||||||||||||      :::|||   ||||||
Db     244 AATCTGAAGGTGAAGAGCGGACAGACGGTGGCCCTGGTTGGCAACAGTGGCTGTGGAAAA 303

Qy     229 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 248
           ||||||   |||||||||:::||||||||||||||||||   ||||||   :::   :::
Db     304 AGCACAACTGTCCAGCTGATGCAAAGGCTCTACGACCCCCTAGATGGCATGGTCAGTATC 363

Qy     249 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 268
           |||   |||||||||   :::||||||||||:::   |||:::   ||||||||||||
Db     364 GACGGACAGGACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCATTGGTGTGGTG 423

Qy     269 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 288
           ||||||||||||||||||||   ||||||||||:::   ||||||:::|||||||||:::
Db     424 AGTCAGGAACCTGTGCTGTTTGCCACCACGATCGCCGAGAACATTCGCTATGGCCGAGAA 483

Qy     289 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 308
           |||||||||   :::|||:::|||:::|||   :::||||||||||||||||||||||||
Db     484 GATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTCATC 543

Qy     309 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 328
           |||:::   |||::::::|||:::||||||||||||||||:::|||||||||:::||||||
Db     544 ATGAAACTGCCCCACCAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGG 603

Qy     329 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 348
           ||||||||||||||||||   ||||||||||||||||||||||||||||||   |||
Db     604 GGACAGAAACAGAGAATCGC-ATTGCCCGGGCCCTGGTCCGCAATCCCAAGTGCCTTTC- 661

Qy     349 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla-Le 368
           :::|||||||||||||||   ||||||:::||||||::::::   |||||||||||| ||
Db     662 GTGGACGAAGCCACCTCAGC-CTGGATACAGAAAGTGAAGCTGGGGTTCAGGCCGCACCT 720

Qy     368 uGluLysAlaSerLysGlyArgThrThrIle------------ValValAlaHisArgLe 384
           |::::::     ||||||  ||||||:::              ::::::     :::||
Db     721 AGATAGGGTACAAAAGGGCNGGACCACATTGGGAAGCTCATCGCTGTTACCGTCGAAGCT 780

Qy     384 uSerThrIleArgSerAlaAspLeuIleVal------ThrLeuLysAspGlyMetLeuAl 402
           |:::   :::|||                :::|||     ||||||||         ::
Db     781 GACGGCCATGCGGCCGAAGGGGGG-CTTGTGGGAACGACTCTGAAAAGC---------TC 830

Qy     402 aGluLysGlyAlaHis 407
           :||||||||   |||
Db     831 TGAAAAAGGTACTCAT 846
```

```
RESULT 11
BG248052
LOCUS       BG248052                1019 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  602359987F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4488404 5',
            mRNA sequence.
ACCESSION   BG248052
VERSION     BG248052.1  GI:12757867
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1019)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10334  row: n  column: 21
            High quality sequence stop: 650.
FEATURES             Location/Qualifiers
     source          1..1019
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4488404"
                     /clone_lib="NCI_CGAP_Mam1"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="10 months, virgin"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
BASE COUNT      276 a    262 c    283 g    196 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              1.9e-77        Length:        1019
Score:                  770.50         Matches:       154
Percent Similarity:     79.17%         Conservative:  36
Best Local Similarity:  64.17%         Mismatches:    47
Query Match:            14.46%         Indels:        3
DB:                     12             Gaps:          1

US-09-873-409-4 (1-1058) x BG248052 (1-1019)
```

```
Qy     813 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 832
           |||||||||::::::|||      |||   |||   ||||||   ||||||   :::
Db       4 GAAGGAAATGTGCAATTTAGTGGAGTCGTGTTCAACTATCCCACCCGACCCAGCATCCCA 63

Qy     833 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 852
           :::|||:::||||||||||||      :::::::::|||:::|||:::|||   |||||||||
Db      64 GTGCTTCAGGGGCTGAGCCTTGAGGTGAAGAAGGGCCAGACGCTGGCCCTGGTGGGCAGC 123

Qy     853 SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGln 872
           ||||||||||||||||||   ||||||||||||:::|||   |||||||||:::
Db     124 AGTGGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAGCGCTTCTACGACCCCATGGCT 183

Qy     873 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 892
           |||   |||      ||||||   :::   |||:::|||||||||||||||||||:::
Db     184 GGATCAGTGTTTCTAGATGGCAAAGAAATAAAGCAACTGAATGTCCAGTGGCTCCGAGCA 243

Qy     893 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 912
           |||:::   ||||||   |||||||||:::||||||:::||||||||||||||||||
Db     244 CAGCTGGGCATTGTGTCCCAAGAGCCCATTCTCTTTGACTGCAGCATCGCAGAGAACATT 303

Qy     913 AlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla 932
           ||||||||||||||||||||||||||||      :::||||||      ||||||
Db     304 GCCTACGGAGACAACAGCCGGGTCGTGTCTTATGAGGAGATTGTGAGGGCAGCCAAGGAG 363

Qy     933 AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeu 952
           ||||||||||||   ||||||:::   ||||||:::||||||||||||:::||||||
Db     364 GCCAACATCCACCAGTTCATCGACTCGCTACCTGATAAATACAACACCAGAGTAGGAGAC 423

Qy     953 LysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeu 972
           ||||||   ||||||||||||||||||||||||:::||||||||||||||||||:::
Db     424 AAAGGCACTCAGCTGTCGGGTGGGCAGAAGCAGCGCATCGCCATCGCACGCGCCCTCGTC 483
```

```
Qy     973 GlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGlu 992
           ::::::|||  ||||||||||||||||||||||||||||||||||||    :::||||||
Db     484 AGACAGCCTCACATTTTACTTCTGGACGAAGCAACATCAGCTCTGGATACAGAAAGTGAA 543

Qy     993 LysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThr 1012
           ||||||||||||   ||||||||||||||||||   ||||||||||:::|||:::
Db     544 AAGGTTGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACCTGCATTGTGATCGCT 603

Qy    1013 HisArgLeuSerAlaIleGlnAsnAlaAspLeuIle-ValValLeuHis-AsnGlyLysI 1032
           ||||||           ||||||||||||||||| ||||||:::    |||    |||:
Db     604 CACCGCTGTCCACCATCCCAGAACGCGGACTTGATCGGTGGTGATTCAAGAACCGGAAGG 663

Qy    1032 leLys---GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
           ::|||        :::          |||:::::::::       :::   |||||||||
Db     664 TCAAGGGAGCACGGGCACCCCAACAGCAAGTGGTGGCGCAGAAGGGCATCTACTTC 719
```

RESULT 12
BG293345

```
LOCUS       BG293345                  726 bp    mRNA    linear   EST 21-FEB-2001
DEFINITION  602390738F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4502552 5',
            mRNA sequence.
ACCESSION   BG293345
VERSION     BG293345.1  GI:13052943
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 726)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10371  row: 1  column: 09
            High quality sequence stop: 658.
FEATURES             Location/Qualifiers
     source          1..726
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4502552"
                     /clone_lib="NIH_MGC_94"
                     /tissue_type="retina"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 3.3 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      189 a     220 c     190 g     127 t
ORIGIN
```

Alignment Scores:

Pred. No.:	1.11e-75	Length:	726
Score:	753.00	Matches:	152
Percent Similarity:	84.29%	Conservative:	25
Best Local Similarity:	72.38%	Mismatches:	30
Query Match:	14.13%	Indels:	3
DB:	12	Gaps:	0

US-09-873-409-4 (1-1058) x BG293345 (1-726)

```
Qy     847 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 866
           :::|||  ||||||||||||||||||||||||||||||     |||||||||||:::|||
Db       2 CTGGCCCTGGTGGGCAGCAGTGGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAGCGC 61

Qy     867 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 886
              |||||||||:::   |||   ||||||   ||||||    :::||||||:::||||||
Db      62 TTCTATGACCCCATGGCTGGATCAGTGCTCTTAGATGGTCAAGAAGCAAAGAAACTCAAT 121

Qy     887 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 906
           ||||||||||||||:::|||:::     ||||||   |||||||||:::||||||:::|||
Db     122 GTCCAGTGGCTCCGAGCTCAACTGGGCATTGTGTCCCAGGAACCCATTCTCTTTGACTGC 181

Qy     907 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 926
           ||||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db     182 AGCATCGCAGAGAACATCGCCTATGGAGACAACAGCCGGGTCGTGCCTCATGATGAGATT 241

Qy     927 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 946
                 ||||||          ||||||||||||    |||||||||    ||||||:::||||||
Db     242 GTGAGGGCAGCCAAGGAGGCCAACATCCACCCCTTCATCGAGACGCTGCCCCAAAAATAT 301

Qy     947 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 966
           ||||||:::||||||   ||||||   |||||||||||||||||||||||||||:::|||
Db     302 AACACAAGAGTAGGAGACAAGGGGACGCAGCTCTCTGGGGGCCAGAAGCAGAGGATTGCC 361

Qy     967 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 986
           ||||||||||||||||:::::::::|||::::::|||||||||||||||||||||||||||
Db     362 ATCGCCCGAGCCCTCATCAGACAGCCTCGGGTCCTACTGCTGGATGAAGCCACGTCAGCT 421

Qy     987 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 1006
           ||||||   :::||||||||||||||||||   ||||||||||||||||||   ||||||
Db     422 CTGGATACTGAGAGTGAAAAGGTTGTCCAGGAAGCACTGGACAAAGCCAGGGAAGGCCGC 481

Qy    1007 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla-AspLeuIleValVa 1026
           ||||||:::   :::   |||||||||||||       ||||||||||||    ||||||||||
Db     482 ACCTGCATTGCGATCGCTCACCGCCTGTCCACCATCCAGAACGCGGCACTTGATCGTGGT 541

Qy    1026 lLeuHisAsnGlyLysIleLysGluGlnGlyThr-HisGlnGluLeuLeuArgAsnArgA 1046
           |:::    |||||||||:::||||||    ||||||  ||||||:::||||||      :::
Db     542 GATTGAGAACGGCAAGGTCAAGGAGCACGGCACCCCACCAGCAGCTGCTGGCGCAGAAGG 601

Qy    1046 spIleTyrPhe-LysLeuValAsn 1053
              |||||||||  :::|||||||||
Db     602 GCATCTATTTCTCAACTGGTCAAC 625
```

RESULT 13
AZ682350

```
LOCUS       AZ682350                  871 bp    DNA     linear   GSS 14-DEC-2000
DEFINITION  ENTKB16TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   AZ682350
VERSION     AZ682350.1  GI:11819496
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
  ORGANISM  Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 871)
  AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Forward
            Class: shotgun
            High quality sequence start: 16
            High quality sequence stop: 860.
FEATURES             Location/Qualifiers
     source          1..871
                     /organism="Entamoeba histolytica"
```

```
                  /strain="HM1:IMSS"
                  /db_xref="taxon:5759"
                  /clone_lib="Entamoeba histolytica Sheared DNA"
                  /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                  Institute for Genomic Research (TIGR), Rockville, MD.
                  Genomic DNA isolated from broth cultures of E. histolytica
                  using a  method described by Clark and Diamond (Clark,
                  C.G., and Diamond, L.S.  (1993) Entamoeba histolytica: a
                  method for isolate identification.  Exp. Parasitol.
                  77:450.). The DNA was mechanically sheared to give a
                  tight size distribution (~2 kb). The v + i method used for
                  the library  construction is described in detail in Smith,
                  H.O. and Venter, J.C.  (Making small insert libraries for
                  whole genome shotgun sequencing  projects. In Genome
                  Sequencing: A Practical Approach, eds. M. Vaudin  and B.
                  Barell, Oxford University Press, 1999)."
BASE COUNT      349 a    117 c    174 g    231 t
ORIGIN

Alignment Scores:
Pred. No.:                1.72e-75      Length:         871
Score:                    752.50        Matches:        151
Percent Similarity:       71.64%        Conservative:   46
Best Local Similarity:    54.91%        Mismatches:     75
Query Match:              14.12%        Indels:         3
DB:                       17            Gaps:           1

US-09-873-409-4 (1-1058) x AZ682350 (1-871)

Qy    153 AlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsn 172
           |||:::   ||||||||:::::::::|||   ||||||:::    |||   |||||
Db     45 GCTAAAGTTGCCGCATTTAATGTTTATCAAACAATTGATAGAATCCCAGATATTGATTGT 104

Qy    173 PheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnVal 192
              |||     |||     |||              |||    :::   |||::::::|||
Db    105 CAGTCTATTGGAGGTGAATGTCCAACTGAGTGTAATGGAAATATTAGATTTGAGGATGTT 164

Qy    193 SerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIle 212
              |||   ||||||:::|||   |||        :::||||||||||||:::|||   |||
Db    165 CAATTTGTTTATCCAACAAGACTGTCTCATCATGTATTAAAAGGACTTGACCTTGAAATT 224

Qy    213 LysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValVal 232
           |||   |||:::|||:::|||||||||||    :::|||     ||||||||||||    :::
Db    225 AAGAAAGGACAAACAATTGCATTAGTTGGAGCATCAGGATGTGGGAAGTCAACTACTATT 284

Qy    233 GlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAsp 252
           ||||||:::||||||    |||||||||:::   |||    :::    :::|||       |||
Db    285 CAATTAATCCAAAGAAATTATGATCCAAATGGTGGAAGAGTAACATTAGACGGAAAAGAT 344

Qy    253 IleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluPro 272
           ||||||    ||||||:::::      |||:::   ||||||:::|||   |||||||||
Db    345 ATACGAGAGTTGAATATCAAATGGTTAAGAAATCAAATAGGATTAGTTGGACAAGAACCA 404

Qy    273 ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAsp--------- 289
           |||||||||         ||||||       ||||||   |||      :::
Db    405 GTGTTGTTTGCAGGAACAATTCGAGAAAATATTATGCTTGGAGCTAAAGAAGGAGCAACA 464

Qy    290 ValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMet 309
              ::::::|||||||||           |||:::   |||||||||:::|||||||||
Db    465 CCAAGTGAAGAAGAGATGATTGAATGTGCTAAAATGGCAAATGCACATGACTTCATTTCT 524

Qy    310 GluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGly 329
           :::    |||      ::::::|||::::::||||||||||||||||    :::|||||||||
Db    525 AAACTTCCAGAAGGATATGACACAATAATTGGAGAAAAAGGAGCATTATTATCAGGAGGA 584

Qy    330 GlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeu 349
           |||||||||||||||||||||||||||||||:::|||||||||    ||||||:::|||
Db    585 CAAAAACAAAGAATTGCAATTGCACGTGCATTGATTCGAAACCCATCTATTCTTCTTCTT 644

Qy    350 AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGlu 369
           ||||||||||||||||||||||||:::::::|||:::       ||||||    |||||||||
Db    645 GATGAAGCTACATCAGCACTTGATACACAAAGTGAAAAGATTGTACAAGAAGCACTTGAA 704

Qy    370 LysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSer 389
           ||||||||||||||||||||||||:::::|||||||||||||||:::|||:::|||:::
Db    705 AAAGCATCTAAAGGAAGAACAACAATTATTGTAGCACATAGACTAACAACTGTTAGAAAT 764

Qy    390 AlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGlu 409
           ||||||   |||                |||    :::    ||||:::|||    |||    |||
Db    765 GCAGATAAAATATGTGTATTTCATCAAGGAGAAATTATTGAACAAGGAAAACATCAAGAA 824

Qy    410 LeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln 424
           ||||||      :::|||    ||||||    ||||||       |||
Db    825 TTAATGGATTTGAAAGGAACATATTATGGATTAGTCAAAAGACAA 869

RESULT 14
BF796582
LOCUS       BF796582                 944 bp    mRNA    linear   EST 12-JAN-2001
DEFINITION  602258463F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4341710 5',
            mRNA sequence.
ACCESSION   BF796582
VERSION     BF796582.1  GI:12101636
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 944)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis Staudt, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9955  row: n  column: 15
            High quality sequence start: 9
            High quality sequence stop: 669.
FEATURES             Location/Qualifiers
     source          1..944
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4341710"
                     /clone_lib="NIH_MGC_85"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.867 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      265 a    241 c    237 g    201 t
ORIGIN

Alignment Scores:
Pred. No.:                2.94e-75      Length:         944
Score:                    751.00        Matches:        166
Percent Similarity:       74.62%        Conservative:   31
Best Local Similarity:    62.88%        Mismatches:     60
Query Match:              14.09%        Indels:         9
DB:                       12            Gaps:           2

US-09-873-409-4 (1-1058) x BF796582 (1-944)

Qy    779 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 798
           |||:::|||||||||      ||||||||||||||    |||    |||::::::|||
Db     39 TATGCTAAAGCTAAGCTGTCTGCAGCCCACTTATTCATGCTGTTTGAAAGACAACCTCTG 98

Qy    799 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 818
           |||||||||    |||:::||||||    |||||||||          |||||||||:::    |||
Db     99 ATTGACAGCTACAGTGAAGAGGGGCTGAAGCCTGATAAATTTGAAGGAAATATAACATTT 158
```

```
Qy     819 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 838
              ||||||  |||  ||||||  |||  :::|||   :::|||:::||||||||
Db     159 AATGAAGTCGTGTTCAACTATCCCACCCGAGCAAACGTGCCAGTGCTTCAGGGGCTGAGC 218

Qy     839 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 858
           |||   :::::::::|||:::|||:::|||  ||||||||||||||||||||||||
Db     219 CTGGAGGTGAAGAAAGGCCAGACACTAGCCCTGGTGGGCAGCAGTGGCTGTGGGAAGAGC 278

Qy     859 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 878
           |||  |||||||||||:::|||   ||||||||:::   |||   ||||||   |||
Db     279 ACGGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCGGGGACAGTGCTTCTCGAT 338

Qy     879 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 898
           |||   :::||||||:::||||||||||||||||||||:::|||:::   ||||||
Db     339 GGTCAAGAAGCAAAGAAACTCAATGTCCAGTGGCTCAGAGCTCAACTCGGAATCGTGTCT 398

Qy     899 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 918
           ||||||||:::||||||:::||||||||||||||||||||||||||||||||||||
Db     399 CAGGAGCCTATCCTATTTGACTGCAGCATTGCCGAGAATATTGCCTATGGAGACAACAGC 458

Qy     919 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 938
           ||||||||     |||||||||     ||||||  |||||||||||||||   |||
Db     459 CGGGTTGTATCACAGGATGAAATCGTGAGTGCAGCCAAAGCTGCCAACATACATCCTTTC 518

Qy     939 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 958
           ||||||  ||||||  ||||||  |||:::||||||  ||||||  |||||||||
Db     519 ATCGAGACGTTACCCCACAAATATGAAACAAGAGTGGGAGATAAGGGGACTCAGCTCTCA 578

Qy     959 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLys-ProLysIleLe 978
           ||||||  |||||||:::|||||||||||||||||||:::::::::  |||:::|||||
Db     579 GGAGGTCACAAACAGAGGATTGCTATTGCCCGAGCCCTCATCAGACAACCCTCAAATCCT 638

Qy     978 uLeuLeuAspGluAlaThrSerAla-LeuAspAsnAspSerGluLysValVal------G 996
           ||||||||||||||||||||||||  ||||||            |||          :
Db     639 CCTGTTGGATGAAGCTACATCAGCCTCTGGATACCTGACAAGTGAACACGGTTGTCCCAA 698

Qy     996 lnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuS 1016
           ::                |||        ||||||||||||||:::   ||||||
Db     699 GAAGCCCCTGTCACACAGCAGAAGAGCGCCGCACCTGCCTTGTGATTGCTCACCGCTGTC 758

Qy    1016 erAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnG 1036
                 ||||||||||||||  |||||||||          |||::::::|||
Db     759 AAC--ATCCAGAATGCAGACTT-ATAGTGGTGTTCCCAACGGGAGAAGTCAAG------- 808

Qy    1036 lyThrHis 1038
             ||||||
Db     809 --ACACAT 814
```

```
RESULT 15
BH139685
LOCUS       BH139685                886 bp    DNA     linear   GSS 07-AUG-2001
DEFINITION  ENTNA47TR Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   BH139685
VERSION     BH139685.1  GI:15098746
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
  ORGANISM  Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 886)
  AUTHORS   Loftus,B., Wang,Z., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library (2001)
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Reverse
            Class: shotgun
            High quality sequence start: 18
            High quality sequence stop: 851.
FEATURES             Location/Qualifiers
     source          1..886
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a  method described by Clark and Diamond (Clark,
                     C.G., and Diamond, L.S.  (1993) Entamoeba histolytica: a
                     method for isolate identification.  Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + i method used for
                     the library  construction is described in detail in Smith,
                     H.O. and Venter, J.C.  (Making small insert libraries for
                     whole genome shotgun sequencing  projects. In Genome
                     Sequencing: A Practical Approach, eds. M. Vaudin  and B.
                     Barell, Oxford University Press, 1999)."
BASE COUNT      348 a    105 c    176 g    257 t
ORIGIN

Alignment Scores:
Pred. No.:              1.93e-74     Length:          886
Score:                  743.50       Matches:         151
Percent Similarity:     73.43%       Conservative:    48
Best Local Similarity:  55.72%       Mismatches:      71
Query Match:            13.95%       Indels:          2
DB:                     17           Gaps:            1

US-09-873-409-4 (1-1058) x BH139685 (1-886)
```

```
Qy     786 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 805
           |||   :::   :::   ::::::::::::||||||  ||||||      |||:::|||
Db      20 GCAAAGAATACTTATGATGTTATTGATAGAAAACCACT-ATTGATTGTTATAGTGAAGAA 78

Qy     806 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 825
           |||:::      :::         :::|||   :::||||||:::::::::   |||   |||
Db      79 GGTGAAACATTTAATGATGTTAAAGGTGAAATTGAATTTAAAGACATTTGTTTCAGATAT 138

Qy     826 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 845
           |||   ||||||||||         :::|||:::|||:::|||   :::|||:::||||||
Db     139 CCAACAAGACCAGACAATTCTGTCTTGAAAGGTATTTCATTCAAAGTAGAACAAGGAAAA 198

Qy     846 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 865
           |||||||||   ||||||:::||||||||||||||||||||||||||||||||:::::::
Db     199 ACTGTTGCATTAGTAGGAGCATCAGGATGTGGTAAATCAACATCAGTTCAGTTGATTGAA 258

Qy     866 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 885
           |||   |||||||||      |||   ||||||   ||||||   :::   |||:::|||
Db     259 AGATTTTATGATCCAACACATGGAGATGTATTATTAGATGGACATAATATCAAAGATTTG 318

Qy     886 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 905
           |||:::   :::|||||||||||||||   :::|||   ||||||||||||||||||
Db     319 AATATTCATTTCTTAAGAAGTCAAATTGGAATGGTAGGACAAGAACCAGTATTATTTGCT 378

Qy     906 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal---ValProLeuAsp 924
              |||:::   :::||||||      |||                |||   |||      :::
Db     379 GAAAGTGTTATGGATAATATTAGAAGAGGAGTACCTAAAGGAGTTGAAGTAAGTAATGAA 438

Qy     925 GluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGlu 944
           :::|||      ||||||      ||||||   |||   ||||||      :::||||||
Db     439 CAAATTTATGCTGCTGCTAAAATGGCAAATGCACATGACTTTATTTCAGCAATGCCAGAA 498

Qy     945 LysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArg 964
              |||||||||   ||||||   :::|||||||||:::||||||||||||||||||
Db     499 GGATATAACACAATGGTAGGTGATAGAGGTGCACAAATTTCAGGAGGACAAAAACAAAGA 558
```

```
Qy     965 LeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThr 984
           :::||||||||||||||||||::::::   |||||||:::|||||||||||||||||||||
Db     559 ATTGCTATTGCACGTGCATTGATTAGAAATCCAAAAGTGTTATTACTCGATGAAGCTACA 618

Qy     985 SerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThr 1004
           ||||||||||||::::::|||||||||:::||||||   |||||||||||||||
Db     619 TCAGCACTTGATTCAGAAAGTGAAAAGATTGTACAAGATGCACTTGACAAAGCAGCAAAA 678

Qy    1005 GlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIle 1024
           |||||||||   :::|||:::   ||||||||||||   |||||||||||||||   |||
Db     679 GGAAGAACAACAATTGTAATTGCACATAGATTATCAACTATTCAAAATGCAGATCAAATA 738

Qy    1025 ValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsn 1044
              |||:::         |||:::|||   |||:::|||||||||||||||||||||
Db     739 TGTGTTATTATGAGAGGAAGAATTGCAGAAAGAGGAACACATCAAGAGTTATTAGATTTG 798

Qy    1045 ArgAspIleTyrPheLysLeuValAsnAlaGln 1055
           :::      |||:::   |||         |||
Db     799 AAAGGATTTATTATACACTTGCTATGCAACAA 831
```

Search completed: March 31, 2003, 13:50:31
Job time : 7305.24 secs

```
              GenCore version 5.1.4_p5_4578
        Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 30, 2003, 01:24:25 ; Search time 14303 Seconds
                                          (without alignments)
                                          2152.747 Million cell updates/sec

Title:          US-09-873-409-4
Perfect score:  5330
Sequence:       1 MVISLTSKELSAYSKAGAVA..........QELLRNRDIYFKLVNAQSVQ 1058

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters:      4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115418_19228/app_query.fasta_1.7544
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_54887_@runat_27032003_115418_19228 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      GenEmbl:*
                1:  gb_ba:*
                2:  gb_htg:*
                3:  gb_in:*
                4:  gb_om:*
                5:  gb_ov:*
                6:  gb_pat:*
                7:  gb_ph:*
                8:  gb_pl:*
                9:  gb_pr:*
                10:  gb_ro:*
                11:  gb_sts:*
                12:  gb_sy:*
                13:  gb_un:*
                14:  gb_vi:*
                15:  em_ba:*
                16:  em_fun:*
                17:  em_hum:*
                18:  em_in:*
                19:  em_mu:*
                20:  em_om:*
                21:  em_or:*
                22:  em_ov:*
                23:  em_pat:*
                24:  em_ph:*
                25:  em_pl:*
                26:  em_ro:*
                27:  em_sts:*
                28:  em_un:*
                29:  em_vi:*
                30:  em_htg_hum:*
                31:  em_htg_inv:*
                32:  em_htg_other:*
                33:  em_htg_mus:*
                34:  em_htg_pln:*
                35:  em_htg_rod:*
                36:  em_htg_mam:*
                37:  em_htg_vrt:*
                38:  em_sy:*
                39:  em_htgo_hum:*
                40:  em_htgo_mus:*
                41:  em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5328	100.0	3177	6	AX339030	AX339030 Sequence
2	5328	100.0	3702	6	AX339031	AX339031 Sequence
3	5159.5	96.8	3621	6	AX339032	AX339032 Sequence
4	4985.5	93.5	3699	6	AX478104	AX478104 Sequence
5	4561.5	85.6	2856	6	AX339028	AX339028 Sequence
6	3456	64.8	2066	6	AX339027	AX339027 Sequence
7	3137.5	58.9	4296	10	CRUPGPI	M60040 C.griseus P
8	3136.5	58.8	3987	10	CRUPGP1185	M59254 Chinese ham
9	3136.5	58.8	4304	10	CRUPGP1165	M59253 Chinese ham
10	3121.5	58.6	4018	5	GGA9799	AJ009799 Gallus ga
11	3118.5	58.5	3858	4	OAU78609	U78609 Ovis aries
12	3116.5	58.5	4927	10	AF257746	AF257746 Rattus no
13	3115.5	58.5	4279	6	AX105082	AX105082 Sequence
14	3112.5	58.4	4279	6	AX105057	AX105057 Sequence
15	3112.5	58.4	4279	6	AX105078	AX105078 Sequence
16	3112.5	58.4	4279	6	AX105080	AX105080 Sequence
17	3111.5	58.4	4317	4	AF045016	AF045016 Canis fam
18	3111.5	58.4	4317	6	AX105059	AX105059 Sequence
19	3110.5	58.4	4045	12	AF269224	AF269224 Synthetic
20	3105.5	58.3	3934	4	CFA419568	AJ419568 Canis fam
21	3103.5	58.2	4323	10	AF286167	AF286167 Rattus no
22	3100.5	58.2	3860	6	AX322787	AX322787 Sequence
23	3100.5	58.2	3860	6	AX322789	AX322789 Sequence
24	3100.5	58.2	4186	6	AX108654	AX108654 Sequence
25	3100.5	58.2	4195	6	AX108656	AX108656 Sequence
26	3100.5	58.2	4264	6	AR051647	AR051647 Sequence
27	3100.5	58.2	4264	6	AR051650	AR051650 Sequence
28	3100.5	58.2	4646	6	AX336420	AX336420 Sequence
29	3100.5	58.2	4646	6	AX336708	AX336708 Sequence
30	3100.5	58.2	4646	6	AX391099	AX391099 Sequence
31	3100.5	58.2	4646	6	I49610	I49610 Sequence 2
32	3100.5	58.2	4646	9	HUMMDR1	M14758 Homo sapien
33	3100.5	58.2	4669	6	I08557	I08557 Sequence 3
34	3100.5	58.2	6505	6	AR028671	AR028671 Sequence
35	3100.5	58.2	8630	6	AX012320	AX012320 Sequence
36	3100.5	58.2	8630	6	AX012321	AX012321 Sequence
37	3100.5	58.2	9318	6	AR028672	AR028672 Sequence
38	3097.5	58.1	4378	6	E02326	E02326 Multidrug r
39	3091	58.0	3888	10	AY082609	AY082609 Rattus no
40	3090.5	58.0	3682	10	MUSPGLY1A	M24417 Mouse phosp
41	3090.5	58.0	4788	6	AX322793	AX322793 Sequence
42	3090.5	58.0	4924	10	MUSMDR1A	M33581 Mouse P-gly
43	3089	58.0	4280	10	CRUPGPII	M60041 C.griseus P
44	3088.5	57.9	4356	10	MUSMDRAA	M30697 Mouse multi
45	3086	57.9	4233	6	AR123273	AR123273 Sequence

ALIGNMENTS

RESULT 1

```
AX339030
LOCUS       AX339030                3177 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 12 from Patent WO0194400.
ACCESSION   AX339030
VERSION     AX339030.1  GI:18129122
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 12 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..3177
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      970 a    601 c    735 g    870 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              0              Length:          3177
Score:                  5328.00        Matches:         1058
Percent Similarity:     100.00%        Conservative:    0
Best Local Similarity:  100.00%        Mismatches:      0
Query Match:            99.96%         Indels:          0
DB:                     6              Gaps:            0

US-09-873-409-4 (1-1058) x AX339030 (1-3177)

Qy       1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCA 60

Qy      21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAGGAGAAAGAACTT 120

Qy      41 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTTGGTTTTATTTCCCCAGTGGCTA 180

Qy      61 LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGAATCTCAAAGATGCAAAGGATTTT 240

Qy      81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAAT 300

Qy     101 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGA 360

Qy     121 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATT 420

Qy     141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATT 480

Qy     161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCT 540

Qy     181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCA 600

Qy     201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 TCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTG 660

Qy     221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 720

Qy     241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 CCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCAT 780

Qy     261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 TATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGT 840

Qy     281 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 AACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGG 900

Qy     301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGG 960

Qy     321 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTA 1020

Qy     341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGC 1080

Qy     361 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 AAGTCAGCTGTTCAAGCTGCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTA 1140

Qy     381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 GCACACCGACTTTCTACTATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATG 1200

Qy     401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 CTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTT 1260

Qy     421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACT 1320

Qy     441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 GAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATT 1380

Qy     461 AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 GACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTA 1440

Qy     481 LysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 AAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCT 1500

Qy     501 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 GTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATG 1560

Qy     521 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 TTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTC 1620

Qy     541 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 GTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGA 1680

Qy     561 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 580
```

```
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1681 GCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAG 1740

Qy        581 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1741 GATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCC 1800

Qy        601 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1801 ATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAAT 1860

Qy        621 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1861 GCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTC 1920

Qy        641 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1921 CTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATG 1980

Qy        661 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1981 ACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAA 2040

Qy        681 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2041 GCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATG 2100

Qy        701 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2101 TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGA 2160

Qy        721 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2161 AGCTGTTATGCATTCAGCCATGCCTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTT 2220

Qy        741 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2221 GGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACT 2280

Qy        761 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2281 GCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCC 2340

Qy        781 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2341 AAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGAC 2400

Qy        801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2401 AGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAA 2460

Qy        821 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2461 GTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGT 2520

Qy        841 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2521 ATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCT 2580

Qy        861 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2581 GTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTG 2640

Qy        881 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2641 GATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAG 2700

Qy        901 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2701 CCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTG 2760

Qy        921 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2761 GTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAA 2820

Qy        941 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2821 GGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGC 2880

Qy        961 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2881 CAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTG 2940

Qy        981 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2941 GATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGAT 3000

Qy       1001 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3001 AAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAC 3060

Qy       1021 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3061 GCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAG 3120

Qy       1041 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1058
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3121 CTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3174

RESULT 2
AX339031
LOCUS       AX339031               3702 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 13 from Patent WO0194400.
ACCESSION   AX339031
VERSION     AX339031.1  GI:18129123
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 13 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .3702
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1098 a    708 c    856 g   1039 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              0               Length:         3702
Score:                  5328.00         Matches:        1058
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            99.96%          Indels:         0
DB:                     6               Gaps:           0

US-09-873-409-4 (1-1058) x AX339031 (1-3702)

Qy          1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        526 ATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCA 585

Qy         21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        586 GAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAGGAGAAAGAACTT 645

Qy         41 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        646 CAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTTGGTTTTATTTCCCCAGTGGCTA 705

Qy         61 LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80
```

```
Db     706 CTAAGTTGTGTTCTGTTNTTGTAAGGTATACACAGAATCTCAAAGATGCAAAGGATTTT 765

Qy      81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 100
Db     766 GGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAAT 825

Qy     101 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120
Db     826 GGAACCTATGGACTTGCTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGA 885

Qy     121 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 140
Db     886 TATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATT 945

Qy     141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
Db     946 GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATT 1005

Qy     161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180
Db    1006 TTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCT 1065

Qy     181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200
Db    1066 GAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCA 1125

Qy     201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
Db    1126 TCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTG 1185

Qy     221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240
Db    1186 GTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 1245

Qy     241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260
Db    1246 CCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCAT 1305

Qy     261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280
Db    1306 TATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGT 1365

Qy     281 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 300
Db    1366 AACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGG 1425

Qy     301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320
Db    1426 GAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGG 1485

Qy     321 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340
Db    1486 GAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTA 1545

Qy     341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360
Db    1546 GTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGC 1605

Qy     361 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380
Db    1606 AAGTCAGCTGTTCAAGCTGCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTA 1665

Qy     381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400
Db    1666 GCACACCGACTTTCTACTATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATG 1725

Qy     401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420
Db    1726 CTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTT 1785

Qy     421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440
Db    1786 GTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACT 1845

Qy     441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460
Db    1846 GAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATT 1905

Qy     461 AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 480
Db    1906 GACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTA 1965

Qy     481 LysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 500
Db    1966 AAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCT 2025

Qy     501 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 520
Db    2026 GTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATG 2085

Qy     521 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 540
Db    2086 TTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTC 2145

Qy     541 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 560
Db    2146 GTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGA 2205

Qy     561 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 580
Db    2206 GCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAG 2265

Qy     581 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600
Db    2266 GATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCC 2325

Qy     601 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620
Db    2326 ATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAAT 2385

Qy     621 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 640
Db    2386 GCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTC 2445

Qy     641 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 660
Db    2446 CTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATG 2505

Qy     661 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 680
Db    2506 ACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAA 2565

Qy     681 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 700
Db    2566 GCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATG 2625

Qy     701 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720
Db    2626 TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGA 2685

Qy     721 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 740
Db    2686 AGCTGTTATGCATTCAGCCATGCCTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTT 2745

Qy     741 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760
Db    2746 GGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACT 2805

Qy     761 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 780
Db    2806 GCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCC 2865

Qy     781 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 800
Db    2866 AAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGAC 2925
```

```
Qy      801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2926 AGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAA 2985

Qy      821 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2986 GTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGT 3045

Qy      841 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3046 ATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCT 3105

Qy      861 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3106 GTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTG 3165

Qy      881 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3166 GATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAG 3225

Qy      901 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3226 CCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTG 3285

Qy      921 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3286 GTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAA 3345

Qy      941 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3346 GGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGC 3405

Qy      961 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3406 CAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTG 3465

Qy      981 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3466 GATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGAT 3525

Qy     1001 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3526 AAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAC 3585

Qy     1021 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3586 GCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAG 3645

Qy     1041 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1058
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3646 CTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3699
```

```
RESULT 3
AX339032
LOCUS       AX339032                3621 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 14 from Patent WO0194400.
ACCESSION   AX339032
VERSION     AX339032.1  GI:18129124
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 14 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .3621
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1081 a    696 c    842 g   1002 t
ORIGIN

Alignment Scores:
Pred. No.:              0             Length:        3621
Score:                  5159.50       Matches:       1031
Percent Similarity:     97.45%        Conservative:  0
Best Local Similarity:  97.45%        Mismatches:    0
Query Match:            96.80%        Indels:        27
DB:                     6             Gaps:          1

US-09-873-409-4 (1-1058) x AX339032 (1-3621)
```

```
Qy        1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      526 ATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCA 585

Qy       21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      586 GAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAGGAGAAAGAACTT 645

Qy       41 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60
            |||
Db      646 CAA--------------------------------------------------------- 648

Qy       61 LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80
                                    ||||||||||||||||||||||||||||||||||||
Db      649 ------------------------AGGTATACACAGAATCTCAAAGATGCAAAGGATTTT 684

Qy       81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      685 GGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAAT 744

Qy      101 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      745 GGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGA 804

Qy      121 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      805 TATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATT 864

Qy      141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      865 GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATT 924

Qy      161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      925 TTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCT 984

Qy      181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      985 GAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCA 1044

Qy      201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1045 TCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTG 1104

Qy      221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1105 GTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 1164

Qy      241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1165 CCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCAT 1224

Qy      261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1225 TATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGT 1284

Qy      281 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1285 AACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGG 1344
```

```
Qy    301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1345 GAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGG 1404

Qy    321 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1405 GAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTA 1464

Qy    341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1465 GTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGC 1524

Qy    361 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1525 AAGTCAGCTGTTCAAGCTGCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTA 1584

Qy    381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1585 GCACACCGACTTTCTACTATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATG 1644

Qy    401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1645 CTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTT 1704

Qy    421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1705 GTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACT 1764

Qy    441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1765 GAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATT 1824

Qy    461 AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1825 GACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTA 1884

Qy    481 LysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1885 AAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCT 1944

Qy    501 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1945 GTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATG 2004

Qy    521 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2005 TTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTC 2064

Qy    541 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2065 GTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGA 2124

Qy    561 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2125 GCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAG 2184

Qy    581 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2185 GATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCC 2244

Qy    601 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2245 ATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAAT 2304

Qy    621 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2305 GCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTC 2364

Qy    641 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2365 CTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATG 2424

Qy    661 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2425 ACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAA 2484

Qy    681 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2485 GCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATG 2544

Qy    701 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2545 TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGA 2604

Qy    721 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2605 AGCTGTTATGCATTCAGCCATGCCTTTATATATTTGCCTATGCAGCAGGGTTTCGATTT 2664

Qy    741 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2665 GGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACT 2724

Qy    761 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2725 GCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCC 2784

Qy    781 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2785 AAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGAC 2844

Qy    801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2845 AGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAA 2904

Qy    821 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2905 GTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGT 2964

Qy    841 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2965 ATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGAAAAAGCACTTCT 3024

Qy    861 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3025 GTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTG 3084

Qy    881 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3085 GATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAG 3144

Qy    901 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3145 CCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTG 3204

Qy    921 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3205 GTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAA 3264

Qy    941 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3265 GGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGC 3324

Qy    961 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3325 CAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTG 3384

Qy    981 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3385 GATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGAT 3444

Qy   1001 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3445 AAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAC 3504

Qy   1021 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1040
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3505 GCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAG 3564

Qy     1041 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1058
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3565 CTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3618

RESULT 4
AX478104
LOCUS       AX478104                3699 bp    DNA     linear   PAT 12-AUG-2002
DEFINITION  Sequence 34 from Patent WO0240541.
ACCESSION   AX478104
VERSION     AX478104.1  GI:22217064
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Tang,Y.T., Yue,H., Nguyen,D.B., Hafalia,A.J., Elliott,V.S., Lu,Y.,
            Walia,N.K., Yao,M.G., Baughn,M.R., Gandhi,A.R., Ding,L.,
            Sanjanwala,M., Ramkumar,J., Arvizu,C., Gietzen,K.J., Lal,P.G.,
            Azimzai,Y., Khan,F.A., Thangavelu,K., Thornton,M., Lu,D.A.,
            Tribouley,C.M., Warren,B.A., Ison,C.H., Das,D., Raumann,B.E.,
            Policky,J.L. and Kearney,L.
  TITLE     Transporters and ion channels
  JOURNAL   Patent: WO 0240541-A 34 23-MAY-2002;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..3699
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 7472030CB1"
BASE COUNT     1116 a    707 c    860 g   1016 t
ORIGIN

Alignment Scores:
Pred. No.:                  2e-317          Length:         3699
Score:                      4985.50         Matches:        1002
Percent Similarity:         94.80%          Conservative:   1
Best Local Similarity:      94.71%          Mismatches:     2
Query Match:                93.54%          Indels:         53
DB:                         6               Gaps:           2

US-09-873-409-4 (1-1058) x AX478104 (1-3699)

Qy        1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      682 ATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCA 741

Qy       21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      742 GAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAGGAGAAAGAACTT 801

Qy       41 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60
            |||
Db      802 CAA--------------------------------------------------------- 804

Qy       61 LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80
                                    ||||||||||||||||||||||||||||||||||||
Db      805 ------------------------AGGTATACACAGAATCTCAAAGATGCAAAGGATTTT 840

Qy       81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 GGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAAT 900

Qy      101 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 GGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGA 960

Qy      121 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 TATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGTAGTTATTGCATT 1020

Qy      141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATT 1080

Qy      161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180
            ||||||||||||||||||||||||||||||   |||||||||||||||||||||||||||
Db     1081 TTCCAGGTTATTGATAAGAAACCCAGTATAGGTAACTTTTCCACAGCTGGATATAAACCT 1140

Qy      181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1141 GAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGGCCA 1200

Qy      201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1201 TCTATCAAGATTCTGAAAGGTCTGAATCTCGGAATTAAGTCTGGAGAGACAGTCGCCTTG 1260

Qy      221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1261 GTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 1320

Qy      241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1321 CCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCAT 1380

Qy      261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1381 TATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGT 1440

Qy      281 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1441 AACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGG 1500

Qy      301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1501 GAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGG 1560

Qy      321 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1561 GAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTA 1620

Qy      341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1621 GTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGC 1680

Qy      361 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1681 AAGTCAGCTGTTCAAGCTGCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTA 1740

Qy      381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1741 GCACACCGACTTTCTACTATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATG 1800

Qy      401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1801 CTGGCGGAGAAAGGAGCGCATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTT 1860

Qy      421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1861 GTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACT 1920

Qy      441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1921 GAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATT 1980

Qy      461 AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1981 GACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTA 2040

Qy      481 LysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2041 AAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCT 2100
```

```
Qy      501 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2101 GTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATG 2160

Qy      521 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2161 TTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTC 2220

Qy      541 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 560
            |||||||||||||||||||||||||||||||||||||||
Db     2221 GTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATG--------------------- 2259

Qy      561 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 580
                                                                     |||
Db     2260 ---------------------------------------------------------CAG 2262

Qy      581 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2263 GATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCC 2322

Qy      601 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2323 ATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAAT 2382

Qy      621 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 640
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2383 GCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTC 2442

Qy      641 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2443 CTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATG 2502

Qy      661 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 680
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2503 ACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAA 2562

Qy      681 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 700
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2563 GCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATG 2622

Qy      701 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2623 TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGA 2682

Qy      721 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2683 AGCTGTTATGCATTCAGCCATGCCTTTATATATTTTGCCTATGCGGCAGGGTTTCGATTT 2742

Qy      741 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2743 GGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACT 2802

Qy      761 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 780
            ||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db     2803 GCAATTGCATATGGAGCTATGGCCATCGGAGAAACGCTCGTTTTGGCTCCTGAATATTCC 2862

Qy      781 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2863 AAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGAC 2922

Qy      801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2923 AGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAA 2982

Qy      821 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2983 GTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGT 3042

Qy      841 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3043 ATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCT 3102

Qy      861 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3103 GTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTG 3162

Qy      881 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3163 GATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAG 3222

Qy      901 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3223 CCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTG 3282

Qy      921 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3283 GTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAA 3342

Qy      941 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3343 GGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGC 3402

Qy      961 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3403 CAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTG 3462

Qy      981 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3463 GATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGAT 3522

Qy     1001 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3523 AAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAC 3582

Qy     1021 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3583 GCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAG 3642

Qy     1041 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1058
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3643 CTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3696
```

```
RESULT 5
AX339028
LOCUS       AX339028                2856 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 10 from Patent WO0194400.
ACCESSION   AX339028
VERSION     AX339028.1  GI:18129120
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 10 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..2856
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      888 a    540 c    652 g    776 t
ORIGIN

Alignment Scores:
Pred. No.:              7.66e-290      Length:          2856
Score:                  4561.50        Matches:         912
Percent Similarity:     97.96%         Conservative:    0
Best Local Similarity:  97.96%         Mismatches:      1
Query Match:            85.58%         Indels:          18
DB:                     6              Gaps:            1

US-09-873-409-4 (1-1058) x AX339028 (1-2856)
```

```
Qy     129 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHis-PheGl 148
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     114 GTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATTGGAGCAGCAGTCCCTCATTATT-- 171

Qy     148 uThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPr 168
                                                             ||||||||||
Db     172 -------------------------------------------------GATAAGAAACC 182

Qy     168 oSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGl 188
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 CAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGA 242

Qy     188 uPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLe 208
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 ATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCT 302

Qy     208 uAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLy 228
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 GAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAA 362

Qy     228 sSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVa 248
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     363 GAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGT 422

Qy     248 lAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVa 268
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     423 GGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTGGT 482

Qy     268 lSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAs 288
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     483 TAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGAGA 542

Qy     288 pAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIl 308
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     543 TGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTTAT 602

Qy     308 eMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGl 328
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     603 CATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGG 662

Qy     328 yGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIl 348
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     663 AGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGAT 722

Qy     348 eLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLe 368
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     723 TTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACT 782

Qy     368 uGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleAr 388
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     783 GGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTAGCACACCGACTTTCTACTATTCG 842

Qy     388 gSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAl 408
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     843 AAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATGCTGGCGGAGAAAGGAGCACATGC 902

Qy     408 aGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLy 428
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     903 TGAACTAATGGCAAAACGAGGTCTATATTATTCACTTGTGATGTCACAGGATATTAAAAA 962

Qy     428 sAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPr 448
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     963 AGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCACTTCC 1022

Qy     448 oLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGl 468
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1023 TCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCCACCCA 1082

Qy     468 nSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPr 488
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1083 ATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAACAAGCC 1142

Qy     488 oGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPr 508
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1143 TGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCTGTTCTAAATGGAACTGTTCATCC 1202

Qy     508 oValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrTh 528
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1203 AGTATTTTCCATCATCTTTGCAAAAATTATAACCATGTTTGGAAATAATGATAAAACCAC 1262

Qy     528 rLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPh 548
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1263 ATTAAAGCATGATGCAGAAATTTATTCCATGATATTCGTCATTTTGGGTGTTATTTGCTT 1322

Qy     548 eValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetAr 568
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1323 TGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGAGCAGGGGAAATTTTAACGATGAG 1382

Qy     568 gLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLy 588
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1383 ATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAGGATATTGCCTGGTTTGATGAAAA 1442

Qy     588 sGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGl 608
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1443 GGAAAACAGCACAGGAGGCTTGACAACAATATTAGCCATAGATATAGCACAAATTCAAGG 1502

Qy     608 yAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVa 628
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1503 AGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAATGCAACTAACATGGGACTTTCAGT 1562

Qy     628 lIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVa 648
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1563 TATCATTTCCTTTATATATGGATGGGAGATGACATTCCTGATTCTGAGTATTGCTCCAGT 1622

Qy     648 lLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLy 668
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1623 ACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATGACTGGATTTGCCAACAAAGATAA 1682

Qy     668 sGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIl 688
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1683 GCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAGAATATACGTACTAT 1742

Qy     688 eValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGl 708
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1743 AGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATGTATGAAGAGATGCTTCAGACTCA 1802

Qy     708 nHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAl 728
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1803 ACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTATGCATTCAGCCATGC 1862

Qy     728 aPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGl 748
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1863 CTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTTGGAGCCTATTTAATTCAAGCTGG 1922

Qy     748 yArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAl 768
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1923 ACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACTGCAATTGCATATGGAGCTATGGC 1982

Qy     768 aIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHi 788
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1983 CATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCCAAAGCCAAATCGGGGGCTGCGCA 2042

Qy     788 sLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLy 808
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2043 TCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGACAGCCGCAGTCAAGAAGGGAAAAA 2102

Qy     808 sProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysAr 828
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2103 GCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTCTTCTATCCATGTCG 2162

Qy     828 gProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAl 848
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2163 CCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGC 2222

Qy     848 aPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTy 868
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    2223 ATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTA 2282

Qy     868 rAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGl 888
           ||||||||||||||.|||||||||||||||||||||||||||||||||||||||||||||
Db    2283 TGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTGGATGCAAAAGAATTGAATGTACA 2342

Qy     888 nTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIl 908
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2343 GTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAGCCTGTGCTCTTCAACTGCAGCAT 2402

Qy     908 eAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGl 928
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2403 TGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTGGTGCCATTAGATGAGATCAAAGA 2462

Qy     928 uAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnTh 948
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2463 AGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAAGGTCTCCCTGAGAAATACAACAC 2522

Qy     948 rGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAl 968
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2523 ACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCTATTGC 2582

Qy     968 aArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAs 988
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2583 AAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGA 2642

Qy     988 pAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCy 1008
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2643 TAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGATAAAGCCAGGACGGGAAGGACATG 2702

Qy    1008 sLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHi 1028
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2703 CCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGGTTCTGCA 2762

Qy    1028 sAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTy 1048
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2763 CAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAGCTCCTGAGAAATCGAGACATATA 2822

Qy    1048 rPheLysLeuValAsnAlaGlnSerValGln 1058
           |||||||||||||||||||||||||||||||
Db    2823 TTTTAAGTTAGTGAATGCACAGTCAGTGCAG 2853

RESULT 6
AX339027
LOCUS       AX339027                2066 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 9 from Patent WO0194400.
ACCESSION   AX339027
VERSION     AX339027.1  GI:18129119
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 9 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..2066
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      647 a    403 c    470 g    546 t
ORIGIN

Alignment Scores:
Pred. No.:              1.42e-217     Length:          2066
Score:                  3456.00       Matches:         687
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            64.84%        Indels:          0
DB:                     6             Gaps:            0

US-09-873-409-4 (1-1058) x AX339027 (1-2066)

Qy     372 SerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAsp 391
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3 AGCAAAGGTCGGACTACAATCGTGGTAGCACACCGACTTTCTACTATTCGAAGTGCAGAT 62

Qy     392 LeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMet 411
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63 TTGATTGTGACCCTAAAGGATGGAATGCTGGCGGAGAAAGGAGCACATGCTGAACTAATG 122

Qy     412 AlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGlu 431
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 GCAAAACGAGGTCTATATTATTCACTTGTGATGTCACAGGATATTAAAAAAGCTGATGAA 182

Qy     432 GlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSer 451
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 CAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCACTTCCTCTGCACTCT 242

Qy     452 ValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGlu 471
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 GTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCCACCCAATCTAAAGAG 302

Qy     472 IleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpPro 491
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 ATAAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAACAAGCCTGAATGGCCT 362

Qy     492 PheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSer 511
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     363 TTTGTGGTTCTGGGGACATTGGCTTCTGTTCTAAATGGAACTGTTCATCCAGTATTTTCC 422

Qy     512 IleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHis 531
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     423 ATCATCTTTGCAAAAATTATAACCATGTTTGGAAATAATGATAAAACCACATTAAAGCAT 482

Qy     532 AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr 551
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     483 GATGCAGAAATTTATTCCATGATATTCGTCATTTTGGGTGTTATTTGCTTTGTCAGTTAT 542

Qy     552 PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHis 571
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     543 TTCATGCAGGGATTATTTTACGGCAGAGCAGGGGAAATTTTAACGATGAGATTAAGACAC 602

Qy     572 LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSer 591
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     603 TTGGCCTTCAAAGCCATGTTATATCAGGATATTGCCTGGTTTGATGAAAAGGAAAACAGC 662

Qy     592 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 611
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     663 ACAGGAGGCTTGACAACAATATTAGCCATAGATATAGCACAAATTCAAGGAGCAACAGGT 722

Qy     612 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 631
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     723 TCCAGGATTGGCGTCTTAACACAAAATGCAACTAACATGGGACTTTCAGTTATCATTTCC 782

Qy     632 PheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 651
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     783 TTTATATATGGATGGGAGATGACATTCCTGATTCTGAGTATTGCTCCAGTACTTGCCGTG 842

Qy     652 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 671
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     843 ACAGGAATGATTGAAACCGCAGCAATGACTGGATTTGCCAACAAAGATAAGCAAGAACTT 902

Qy     672 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 691
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     903 AAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAGAATATACGTACTATAGTGTCATTA 962

Qy     692 ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn 711
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     963 ACAAGGGAAAAAGCCTTCGAGCAAATGTATGAAGAGATGCTTCAGACTCAACACAGAAAT 1022

Qy     712 ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr 731
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    1023 ACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTATGCATTCAGCCATGCCTTTATATAT 1082

Qy     732 PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr 751
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1083 TTTGCCTATGCAGCAGGGTTTCGATTTGGAGCCTATTTAATTCAAGCTGGACGAATGACC 1142

Qy     752 ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys 771
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1143 CCAGAGGGCATGTTCATAGTTTTTACTGCAATTGCATATGGAGCTATGGCCATCGGAAAA 1202

Qy     772 ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAla 791
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1203 ACGCTCGTTTTGGCTCCTGAATATTCCAAAGCCAAATCGGGGGCTGCGCATCTGTTTGCC 1262

Qy     792 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 811
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1263 TTGTTGGAAAAGAAACCAAATATAGACAGCCGCAGTCAAGAAGGGAAAAAGCCAGACACA 1322

Qy     812 CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspVal 831
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1323 TGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTCTTCTATCCATGTCGCCCAGATGTT 1382

Qy     832 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 851
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1383 TTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGCATTTGTGGGG 1442

Qy     852 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProVal 871
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1443 AGCAGCGGCTGTGGGAAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTG 1502

Qy     872 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArg 891
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1503 CAAGGACAAGTGCTGTTTGATGGTGTGGATGCAAAAGAATTGAATGTACAGTGGCTCCGT 1562

Qy     892 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 911
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1563 TCCCAAATAGCAATCGTTCCTCAAGAGCCTGTGCTCTTCAACTGCAGCATTGCTGAGAAC 1622

Qy     912 IleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn 931
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1623 ATCGCCTATGGTGACAACAGCCGTGTGGTGCCATTAGATGAGATCAAAGAAGCCGCAAAT 1682

Qy     932 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 951
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1683 GCAGCAAATATCCATTCTTTTATTGAAGGTCTCCCTGAGAAATACAACACACAAGTTGGA 1742

Qy     952 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 971
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1743 CTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTT 1802

Qy     972 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 991
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1803 CTCCAAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGATAATGACAGT 1862

Qy     992 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 1011
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1863 GAGAAGGTGGTTCAGCATGCCCTTGATAAAGCCAGGACGGGAAGGACATGCCTAGTGGTC 1922

Qy    1012 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 1031
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1923 ACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGGTTCTGCACAATGGAAAG 1982

Qy    1032 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 1051
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1983 ATAAAGGAACAAGGAACTCATCAAGAGCTCCTGAGAAATCGAGACATATATTTTAAGTTA 2042

Qy    1052 ValAsnAlaGlnSerValGln 1058
           |||||||||||||||||||||
Db    2043 GTGAATGCACAGTCAGTGCAG 2063

RESULT 7
CRUPGPI
LOCUS       CRUPGPI                 4296 bp    mRNA    linear   ROD 22-JUN-1995
DEFINITION  C.griseus P-glycoprotein (isoform class I) mRNA, complete cds.
ACCESSION   M60040
VERSION     M60040.1  GI:191164
KEYWORDS    multidrug resistance; p-glycoprotein; transmembrane protein.
SOURCE      C.griseus adult liver and ovary, cDNA to mRNA.
  ORGANISM  Cricetulus griseus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
            Cricetulus.
REFERENCE   1  (bases 1 to 4296)
  AUTHORS   Endicott,J.A., Sarangi,F. and Ling,V.
  TITLE     Complete cDNA sequences encoding the Chinese hamster P-glycoprotein
            gene family
  JOURNAL   DNA Seq. 2 (2), 89-101 (1991)
  MEDLINE   92135896
   PUBMED   1685679
FEATURES             Location/Qualifiers
     source          1..4296
                     /organism="Cricetulus griseus"
                     /db_xref="taxon:10029"
                     /tissue_type="liver and ovary"
                     /dev_stage="adult"
     CDS             103..3933
                     /note="putative"
                     /codon_start=1
                     /product="p-glycoprotein isoform I"
                     /protein_id="AAA68883.1"
                     /db_xref="GI:191165"
                     /translation="MEFEEDFSGRKDKNFLKMGRKSKKEKKEKKPVVSVFTMFRYAGW
                     LDRLYMLVGTLAAIIHGVALPLMMLVFGDMTDSFASVGNIPTNATNNATQVNASDIFG
                     KLEEEMTTYAYYYTGIGAGVLIVAYIQVSFWCLAAGRQIHKIRQKFFHAIMNQEIGWF
                     DVHDVGELNTRLTDDVSKINEGIGDKIGMFFQAMATFFGGFIIGFTRGWKLTLVILAI
                     SPVLGLSAGIWAKILSSFTDKELQAYAKAGAVAEEVLAAIRTVIAFGGQKKELERYNN
                     NLEEAKRLGIKKAITANISMGAAFLLIYASYALAFWYGTSLVISKEYSIGQVLTVFFA
                     VLIGAFSIGQASPNIEAFANARGAAYEIFNIIDNKPSIDSFSKNGYKPDNIKGNLEFK
                     NIHFSYPSRKDVQILKGLNLKVQSGQTVALVGNSGCGKSTTVQLLQRLYDPTEGVVSI
                     DGQDIRTINVRYLREIIGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKEANAYD
                     FIMKLPHKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTESEAVV
                     QAALDKAREGRTTIVIAHRLSTVRNADIIAGFDGGVIVEQGNHEELMREKGIYFKLVM
                     TQTAGNEIELGNEVGESKNEIDNLDMSSKDSASSLIRRRSTRRSIRGPHDQDRKLSTK
                     EALDEDVPPISFWRILKLNSSEWPYFVVGIFCAIVNGALQPAFSIIFSKVVGVFTRNT
                     DDETKRHDSNLFSLLFLILGVISFITFFLQGFTFGKAGEILTKRLRYMVFKSMLRQDV
                     SWFDNPKNTTGALTTRLANDAGQVKGATGARLAVITQNIANLGTGIIISLIYGWQLTL
                     LLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGSGKIATEAIENFRTVVSLTREQKFE
                     NMYAQSLQIPYRNALKKAHVFGITFSFTQAMMYFSYAACFRFGAYLVARELMTFENVL
                     LVFSAIVFGAMAVGQVSSFAPDYAKAKVSASHIIMIIEKVPSIDSYSTGGLKPNTLEG
                     NVKFNEVVFNYPTRPDIPVLQGLNLEVKKGQTLALVGSSGCGKSTVVQLLERFYDPMA
                     GTVFLDGKEVNQLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSQDEIERAA
                     KEANIHQFIESLPDKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEATSALD
                     TESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVIQNGKVKEHGTHQQLLAQKG
                     IYFSMVSVQAGAKR"
BASE COUNT     1271 a    892 c   1029 g   1104 t
ORIGIN

Alignment Scores:
Pred. No.:              2.61e-196      Length:        4296
Score:                  3137.50        Matches:       598
Percent Similarity:     75.82%         Conservative:  211
Best Local Similarity:  56.04%         Mismatches:    215
Query Match:            58.86%         Indels:        43
DB:                     10             Gaps:          6

US-09-873-409-4 (1-1058) x CRUPGPI (1-4296)

Qy       4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
           |||   |||   |||||||||   ||||||:::|||||||||||||||||||||||||||
Db     805 TCATTTACTGATAAAGAACTTCAGGCATATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTC 864

Qy      24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
           |||::::::|||||||||||||||||||||      |||:::|||||||||:::
Db     865 TTAGCAGCCATCAGAACTGTGATTGCATTTGGAGGACAAAAGAAAGAACTTGAA------ 918

Qy      44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
```

```
Db    918 ------------------------------------------------------------ 918

Qy     64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
                         ||||||     ||||||:::::||||||     |||||||||
Db    919 ---------------AGGTACAACAACAATTTAGAAGAAGCCAAAAGACTTGGGATAAAG 963

Qy     84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
          :::   |||   :::   ::|||::|||||   :::       :::      :::|||
Db    964 AAAGCGATCACAGCCAACATTTCCATGGGCGCAGCTTTCCTTCTTATCTATGCATCATAT 1023

Qy    104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
             ||||||||||||||||||||||||||||:::::::::   |||      |||:::|||
Db   1024 GCTCTGGCCTTCTGGTATGGAACGTCCTTGGTCATCTCAAAAGAA------TATTCTATT 1077

Qy    124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
          |||   ||||||   |||||||||:::|||:::      ::::::   ||||||   |||
Db   1078 GGACAAGTGCTCACTGTCTTCTTTGCTGTATTAATTGGGGCATTCAGTATTGGACAGGCA 1137

Qy    144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
             |||:::   |||   ||||||   |||||||||||||||:::   ||||||   :::
Db   1138 TCTCCAAATATCGAAGCCTTTGCCAATGCAAGAGGAGCAGCCTATGAAATCTTCAACATA 1197

Qy    164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
          ||||||   ||||||||||||||||||:::|||||||     ||||||||||||:::::|||
Db   1198 ATTGATAATAAGCCCAGTATTGACAGCTTCTCAAAGAATGGGTACAAACCAGACAACATT 1257

Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
          :::|||   :::||||||||||||:::   |||:::||||||||||||      ::::::
Db   1258 AAAGGAAATTTGGAATTCAAAAATATTCACTTCAGTTACCCATCTCGAAAAGACGTTCAG 1317

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
          ||||||||||||||||||||||||:::::::::||||||:::||||||||||||||||
Db   1318 ATCTTGAAGGGCCTCAACCTGAAGGTGCAGAGCGGACAGACAGTGGCCCTGGTTGGCAAC 1377

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
          :::|||   ||||||||||||   |||||||||||||||||||||||||||   :::
Db   1378 AGTGGCTGTGGGAAAAGCACCACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCACAGAG 1437

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
          |||   :::   :::|||   |||   :::||||||||||:::   |||:::
Db   1438 GGCGTGGTCAGTATCGACGGTCAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAA 1497

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
             |||||||||||||||||||||||||||||   |||||||:::   ||||||
Db   1498 ATCATTGGGGTCGTGAGTCAGGAACCTGTGTTGTTTGCCACCACTATCGCTGAAAACATT 1557

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
          :::||||||||:::::::||||||   :::|||:::|||:::|||   :::|||||||||
Db   1558 CGCTATGGCCGAGAAAATGTCACTATGGACGAGATCGAGAAAGCTGTCAAGGAAGCCAAT 1617

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
          ||||||||||||||||||:::   |||:::||||||:::||||||||||||||:::|||
Db   1618 GCCTATGACTTCATCATGAAACTGCCCCATAAATTTGACACTCTTGTTGGTGAGAGAGGA 1677

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
          |||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1678 GCACAGCTGAGTGGAGGACAGAAACAGAGAATCGCCATTGCTCGTGCCCTGGTCCGCAAC 1737

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
          ||||||||||||||||||||||||||||||||||||||||||:::||||||:::::
Db   1738 CCCAAGATCCTTTTGTTGGATGAGGCGACATCAGCCTTGGACACAGAAAGTGAAGCCGTG 1797

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
          |||||||||||||||:::||||||   :::||||||||||||||||||:::|||||||||
Db   1798 GTTCAGGCTGCTCTGGATAAGGCTAGAGAAGGCCGGACTACCATTGTGATAGCCCACCGC 1857

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
          |||||||||:::|||:::||||||:::|||          |||::::::   |||
Db   1858 TTGTCTACAGTTCGAAATGCTGACATCATAGCTGGGTTTGATGGTGGTGTCATTGTGGAG 1917

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
          :::|||   |||   |||||||||   ::::::|||:::|||:::   |||||||||:::

Db   1918 CAAGGAAATCATGAAGAGCTCATGAGAGAGAAGGGCATTTACTTCAAACTTGTCATGACA 1977

Qy    424 Gln-----------------------------AspIleLysLysAlaAspGluGlnMet 433
          |||                             ::::::   ::::::      ::::::
Db   1978 CAGACAGCAGGAAATGAAATTGAATTAGGAAATGAAGTTGGTGAGTCTAAAAATGAAATT 2037

Qy    434 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu------ProLeuHisSer 451
          :::::::::      |||:::::::        :::||||||            :::
Db   2038 GATAATTTAGACATGTCTTCAAAAGATTCAGCATCCAGTCTAATCAGAAGGAGATCAACT 2097

Qy    452 ValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLys 470
             :::||||||:::                  :::   ||||||   ::::::
Db   2098 CGCAGAAGTATCCGTGGACCACATGACCAAGACAGGAAGCTTAGTACCAAAGAGGCCTTG 2157

Qy    471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
          :::      :::|||   :::|||      :::|||||||||||||||      ||||||
Db   2158 GATGAAGATGTACCTCCAATTTCCTTTTGGCGTATCCTAAAGTTGAATTCATCTGAATGG 2217

Qy    491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
          |||:::   |||:::|||      :::::::::||||||   :::   |||   |||
Db   2218 CCATATTTTGTGGTTGGTATATTCTGTGCCATAGTAAATGGAGCCTTGCAACCAGCATTC 2277

Qy    511 SerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeu 529
          |||||||||||:::|||:::::::   :::|||      |||   |||      |||
Db   2278 TCAATAATATTTTCTAAGGTTGTAGGGGTTTTCACAAGAAATACTGATGATGAAACCAAA 2337

Qy    530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
          :::||||||:::   ::::::|||:::|||:::|||||||||||||||   |||:::
Db   2338 CGACATGATAGCAACTTGTTTTCCTTATTGTTTCTGATCCTTGGGGTCATTTCTTTTATT 2397

Qy    550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
          ::::::|||:::||||||      :::|||:::||||||||||||||||   ||||||
Db   2398 ACATTTTTCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGACTC 2457

Qy    570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
          |||::::::   ||||||:::||||||   ||||||::::::|||||||||      :::
Db   2458 CGATACATGGTTTTCAAATCCATGCTGAGACAGGATGTGAGCTGGTTTGATAACCCTAAA 2517

Qy    590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
          |||:::||||||   |||||||||   ||||||   |||      |||::::::||||||
Db   2518 AACACCACTGGAGCATTGACCACCAGGCTCGCCAACGATGCTGGTCAAGTTAAAGGGGCT 2577

Qy    610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
          ||||||:::|||:::      |||:::||||||   |||:::|||         :::|||
Db   2578 ACAGGAGCCAGACTTGCTGTCATTACCCAGAATATAGCAAATCTTGGGACAGGAATCATC 2637

Qy    630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
          ||||||   |||||||||||:::::::|||   |||:::|||:::|||   |||::::::
Db   2638 ATATCCCTAATCTATGGCTGGCAGTTGACACTTCTACTCTTAGCAATTGTCCCCATCATT 2697

Qy    650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
          |||:::   |||::::::|||         ::::::|||   |||   |||||||||:::
Db   2698 GCAATAGCAGGAGTGGTTGAGATGAAAATGTTGTCTGGACAAGCACTAAAAGATAAGAAG 2757

Qy    670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
          ||||||:::   :::|||||||||||||||||||||:::||||||   ||||||:::|||
Db   2758 GAGCTAGAAGGTTCTGGGAAGATTGCTACTGAAGCAATAGAGAACTTCCGCACTGTCGTC 2817

Qy    690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
          |||||||||||||:::   ||||||   ||||||   :::   ||||||      :::
Db   2818 TCTTTGACTCGGGAGCAGAAGTTTGAAAATATGTATGCCCAGAGCTTGCAGATACCATAC 2877

Qy    710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
          ||||||   |||||||||   :::   ::::::|||:::   |||
Db   2878 AGAAATGCTCTGAAGAAAGCACACGTCTTTGGGATCACATTCTCCTTCACCCAGGCCATG 2937

Qy    730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
          :::||||||:::|||||||||   |||||||||||||||||||||:::
Db   2938 ATGTATTTTTCGTATGCTGCTTGTTTCCGGTTTGGTGCTTACTTGGTGGCACGGGAATTA 2997

Qy    750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
          ||||||   |||   :::   :::|||||||||:::||||||   :::|||||||||:::
Db   2998 ATGACATTTGAAAATGTTCTATTAGTATTCTCAGCTATTGTCTTTGGTGCCATGGCAGTT 3057
```

```
Qy     770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
           |||:::            ||||||:::|||:::||||||||      |||:::|||:::
Db    3058 GGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTGTCGGCATCGCATATC 3117

Qy     790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
                 ::::::||||||   |||:::|||||||||   |||      |||   ||||||
Db    3118 ATCATGATCATTGAAAAAGTCCCTTCCATTGACAGCTACAGCACCGGAGGCCTGAAGCCT 3177

Qy     810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
           :::|||   |||||||||::::::|||   ||||||   |||   ||||||   ||||||
Db    3178 AATACATTGGAAGGAAATGTGAAATTTAATGAAGTCGTGTTCAACTACCCCACCCGGCCT 3237

Qy     830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
           |||:::   :::|||:::||||||:::|||   :::::::::|||:::|||:::|||
Db    3238 GACATCCCAGTGCTTCAGGGGCTGAACCTGGAGGTGAAGAAGGGCCAGACGCTGGCCCTG 3297

Qy     850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
           |||||||||||||||||||||||||||||||   ||||||||||||:::|||   ||||||
Db    3298 GTGGGCAGCAGTGGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAGCGGTTCTACGAC 3357

Qy     870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
           |||:::   |||   |||      ||||||   :::      :::||||||||||||||||
Db    3358 CCCATGGCCGGAACAGTGTTTCTAGATGGCAAAGAAGTAAACCAACTAAATGTCCAGTGG 3417

Qy     890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
           ||||||:::   :::   ||||||   |||||||||:::||||||:::||||||||||||
Db    3418 CTCCGAGCACACCTAGGCATTGTGTCCCAAGAGCCCATCCTGTTTGACTGCAGCATCGCT 3477

Qy     910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
           |||||||||||||||||||||||||||||||||      |||||||||:::   |||
Db    3478 GAGAACATCGCTTATGGAGACAACAGCCGGGTCGTGTCCCAGGATGAGATCGAGAGGGCG 3537

Qy     930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
           |||   ||||||||||||   |||||||||   ||||||:::||||||||||||:::
Db    3538 GCGAAGGAAGCCAACATCCACCAGTTCATTGAGTCACTGCCTGACAAATACAACACCAGA 3597

Qy     950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
           ||||||   ||||||   |||||||||||||||||||||||||:::||||||||||||
Db    3598 GTAGGAGACAAAGGGACTCAGCTGTCGGGTGGGCAGAAACAACGCATCGCCATAGCACGC 3657

Qy     970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
           ||||||:::::::::|||   |||||||||||||||||||||||||||||||||||
Db    3658 GCCCTCGTCAGACAGCCTCACATTTTGCTCCTGGATGAAGCAACATCAGCCCTGGATACA 3717

Qy     990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
           :::||||||||||||||||||   ||||||||||||||||||   ||||||||||||:::
Db    3718 GAGAGTGAAAAGGTTGTCCAGGAAGCTCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 3777

Qy    1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
           |||:::   ||||||||||||   ||||||||||||||||||||||||:::   |||
Db    3778 GTGATCGCTCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATTCAGAAT 3837

Qy    1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
           ||||||:::||||||   ||||||||||||:::||||||      :::   ||||||||
Db    3838 GGCAAGGTCAAGGAGCATGGCACCCACCAGCAGCTGCTGGCACAGAAAGGCATCTATTTC 3897

Qy    1050 LysLeuValAsnAlaGlnSer 1056
              :::|||:::   |||:::
Db    3898 TCCATGGTCAGTGTGCAGGCT 3918
```

RESULT 8
CRUPGP1185

```
LOCUS       CRUPGP1185            3987 bp    mRNA    linear   ROD 27-APR-1993
DEFINITION  Chinese hamster p-glycoprotein mRNA (clone ADX185), complete cds.
ACCESSION   M59254
VERSION     M59254.1  GI:191156
KEYWORDS    multidrug resistance glycoprotein; p-glycoprotein.
SOURCE      Chinese hamster, cDNA to mRNA, clone ADX185.
  ORGANISM  Cricetulus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
            Cricetulus.
REFERENCE   1  (bases 1 to 3987)
  AUTHORS   Devine,S.E., Hussain,A., Davide,J.P. and Melera,P.W.
  TITLE     Full length and alternatively spliced pgp1 transcripts in
            multidrug-resistant Chinese hamster lung cells
  JOURNAL   J. Biol. Chem. 266 (7), 4545-4555 (1991)
  MEDLINE   91154265
   PUBMED   1671863
FEATURES             Location/Qualifiers
     source          1..3987
                     /organism="Cricetulus sp."
                     /db_xref="taxon:10031"
                     /cell_type="fibroblast"
                     /tissue_type="lung"
     gene            1..3987
                     /gene="pgp-1"
     CDS             115..3624
                     /gene="pgp-1"
                     /codon_start=1
                     /product="p-glycoprotein"
                     /protein_id="AAA37005.1"
                     /db_xref="GI:191157"
                     /translation="MTTYAYYYTGIGAGVLIVAYIQVSFWCLAAGRQIHKIRQKFFHA
                     IMNQEIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQAMATFFGGFIIGFTRGW
                     KLTLVILAISPVLGLSAGIWAKILSSFTDKELQAYAKAGAVAEEVLAAIRTVIAFGGQ
                     KKELERYNNNLEEAKRLGIKKAITANISMGAAFLLIYASYALAFWYGTSLVISKEYSI
                     GQVLTVFFAVLIAPFSIGQASPNIEAFANARGAAYEIFNIIDNKPSIDSFSKNGYKPD
                     NIKGNLEFKNIHFSYPSRKDVQILKGLNLKVQSGQTVALVGNSGCGKSTTVQLLQRLY
                     DPTEGVVSIDGQDIRTINVRYLREIIGVVSQEPVLFATTIAENIRYGRENVTMDEIEK
                     AVKEANAYDFIMKLPHKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSA
                     LDTESEAVVQAALDKAREGRTTIVIAHRLSTVRNADIIAGFDGGVIVEQGNHEELMRE
                     KGIYFKLVMTQTAGNEIELGNEVGESKNEIDNLDMSSKDSASSLIRRRSTRRSIRGPH
                     DQDRKLSTKEALDEDVPPISFWRILKLNSSEWPYFVVGIFCAIVNGALQPAFSIIFSK
                     VVGVFTRNTDDETKRHDSNLFSLLFLILGVISFITFFLQGFTFGKAGEILTKRLRYMV
                     FKSMLRQDVSWFDNPKNTTGALTTRLANDAGQVKGATGARLAVITQNIANLGTGIIIS
                     LIYGWQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGSGKIATEAIENFRTVV
                     SLTREQKFENMYAQSLQIPYRNALKKAHVFGITFSFTQAMMYFSYAACFRFGAYLVAR
                     ELMTFENVLLVFSAIVFGAMAVGQVSSFAPDYAKAKVSASHIIMIIEKVPSIDSYSTG
                     GLKPNTLEGNVKFNEVVFNYPTRPDIPVLQGLNLEVKKGQTLALVGSSGCGKSTVVQL
                     LERFYDPMAGTVFLDGKEVNQLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVV
                     SQDEIERAAKEANIHQFIESLPDKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILL
                     LDEATSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVIQNGKVKEHGT
                     HQQLLAQKGIYFSMVSVQAGAKR"
BASE COUNT     1194 a    822 c    944 g   1027 t
ORIGIN      Map position 1q26.
```

```
Alignment Scores:
Pred. No.:              2.76e-196     Length:        3987
Score:                  3136.50       Matches:       598
Percent Similarity:     75.82%        Conservative:  211
Best Local Similarity:  56.04%        Mismatches:    215
Query Match:            58.85%        Indels:        43
DB:                     10            Gaps:          6
```

US-09-873-409-4 (1-1058) x CRUPGP1185 (1-3987)

```
Qy       4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
              |||   |||   |||||||||      ||||||:::||||||||||||||||||||
Db     496 TCATTTACTGATAAAGAACTTCAGGCATATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTC 555

Qy      24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
           |||::::::||||||||||||||||||||||      |||:::|||||||||:::
Db     556 TTAGCAGCCATCAGAACTGTGATTGCATTTGGAGGACAAAAGAAAGAACTTGAA------ 609

Qy      44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63

Db     609 ------------------------------------------------------------ 609

Qy      64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
                        ||||||      ||||||:::::::||||||      |||||||||
Db     610 ---------------AGGTACAACAACAATTTAGAAGAAGCCAAAAGACTTGGGATAAAG 654

Qy      84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
           :::   |||   :::   :::|||:::||||||   :::      :::      :::|||
```

```
Db    655 AAAGCGATCACAGCCAACATTTCCATGGGCGCAGCTTTCCTTCTTATCTATGCATCATAT 714

Qy    104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db    715 GCTCTGGCCTTCTGGTATGGAACGTCCTTGGTCATCTCAAAAGAA------TATTCTATT 768

Qy    124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db    769 GGACAAGTGCTCACTGTCTTCTTTGCTGTATTAATTGCGCCATTCAGTATTGGACAGGCA 828

Qy    144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db    829 TCTCCAAATATCGAAGCCTTTGCCAATGCAAGAGGAGCAGCCTATGAAATCTTCAACATA 888

Qy    164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db    889 ATTGATAATAAGCCCAGTATTGACAGCTTCTCAAAGAATGGGTACAAACCAGACAACATT 948

Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db    949 AAAGGAAATTTGGAATTCAAAAATATTCACTTCAGTTACCCATCTCGAAAAGACGTTCAG 1008

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db   1009 ATCTTGAAGGGCCTCAACCTGAAGGTGCAGAGCGGACAGACAGTGGCCCTGGTTGGCAAC 1068

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db   1069 AGTGGCTGTGGGAAAAGCACCACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCACAGAG 1128

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db   1129 GGCGTGGTCAGTATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAA 1188

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db   1189 ATCATTGGGGTCGTGAGTCAGGAACCTGTGTTGTTTGCCACCACTATCGCTGAAAACATT 1248

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db   1249 CGCTATGGCCGAGAAAATGTCACTATGGACGAGATCGAGAAAGCTGTCAAGGAAGCCAAT 1308

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db   1309 GCCTATGACTTCATCATGAAACTGCCCCATAAATTTGACACTCTTGTTGGTGAGAGAGGA 1368

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db   1369 GCACAGCTGAGTGGAGGACAGAAACAGAGAATCGCCATTGCTCGTGCCCTGGTCCGCAAC 1428

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db   1429 CCCAAGATCCTTTTGTTGGATGAGGCGACATCAGCCTTGGACACAGAAAGTGAAGCCGTG 1488

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db   1489 GTTCAGGCTGCTCTGGATAAGGCTAGAGAAGGCCGGACTACCATTGTGATAGCCCACCGC 1548

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db   1549 TTGTCTACAGTTCGAAATGCTGACATCATAGCTGGGTTTGATGGTGGTGTCATTGTGGAG 1608

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   1609 CAAGGAAATCATGAAGAGCTCATGAGAGAGAAGGGCATTTACTTCAAACTTGTCATGACA 1668

Qy    424 Gln------------------------------AspIleLysLysAlaAspGluGlnMet 433
Db   1669 CAGACAGCAGGAAATGAAATTGAATTAGGAAATGAAGTTGGTGAGTCTAAAAATGAAATT 1728

Qy    434 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu------ProLeuHisSer 451
Db   1729 GATAATTTAGACATGTCTTCAAAAGATTCAGCATCCAGTCTAATCAGAAGGAGATCAACT 1788

Qy    452 ValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLys 470
Db   1789 CGCAGAAGTATCCGTGGACCACATGACCAAGACAGGAAGCTTAGTACCAAAGAGGCCTTG 1848

Qy    471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
Db   1849 GATGAAGATGTACCTCCAATTTCCTTTTGGCGTATCCTAAAGTTGAATTCATCTGAATGG 1908

Qy    491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db   1909 CCATATTTTGTGGTTGGTATATTCTGTGCCATAGTAAATGGAGCCTTGCAACCAGCATTC 1968

Qy    511 SerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeu 529
Db   1969 TCAATAATATTTTCTAAGGTTGTAGGGGTTTTCACAAGAAATACTGATGATGAAACCAAA 2028

Qy    530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db   2029 CGACATGATAGCAACTTGTTTTCCTTATTGTTTCTGATCCTTGGGGTCATTTCTTTTATT 2088

Qy    550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db   2089 ACATTTTTTCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGACTC 2148

Qy    570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
Db   2149 CGATACATGGTTTTCAAATCCATGCTGAGACAGGATGTGAGCTGGTTTGATAACCCTAAA 2208

Qy    590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db   2209 AACACCACTGGAGCATTGACCACCAGGCTCGCCAACGATGCTGGTCAAGTTAAAGGGGCT 2268

Qy    610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db   2269 ACAGGAGCCAGACTTGCTGTCATTACCCAGAATATAGCAAATCTTGGGACAGGAATCATC 2328

Qy    630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db   2329 ATATCCCTAATCTATGGCTGGCAGTTGACACTTCTACTCTTAGCAATTGTCCCCATCATT 2388

Qy    650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db   2389 GCAATAGCAGGAGTGGTTGAGATGAAAATGTTGTCTGGACAAGCACTAAAAGATAAGAAG 2448

Qy    670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db   2449 GAGCTAGAAGGTTCTGGGAAGATTGCTACTGAAGCAATAGAGAACTTCCGCACTGTCGTC 2508

Qy    690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db   2509 TCTTTGACTCGGGAGCAGAAGTTTGAAAATATGTATGCCCAGAGCTTGCAGATACCATAC 2568

Qy    710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db   2569 AGAAATGCTCTGAAGAAAGCACACGTCTTTGGGATCACATTCTCCTTCACCCAGGCCATG 2628

Qy    730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db   2629 ATGTATTTTTCGTATGCTGCTTGTTTCCGGTTTGGTGCTTACTTGGTGGCACGGGAATTA 2688

Qy    750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db   2689 ATGACATTTGAAAATGTTCTATTAGTATTCTCAGCTATTGTCTTTGGTGCCATGGCAGTT 2748

Qy    770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db   2749 GGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTGTCGGCATCGCATATC 2808

Qy    790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db   2809 ATCATGATCATTGAAAAAGTCCCTTCCATTGACAGCTACAGCACCGGAGGCCTGAAGCCT 2868
```

```
Qy    810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db   2869 AATACATTGGAAGGAAATGTGAAATTTAATGAAGTCGTGTTCAACTACCCCACCCGGCCT 2928

Qy    830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db   2929 GACATCCCAGTGCTTCAGGGGCTGAACCTGGAGGTGAAGAAGGGCCAGACGCTGGCCCTG 2988

Qy    850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db   2989 GTGGGCAGCAGTGGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAGCGGTTCTACGAC 3048

Qy    870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
Db   3049 CCCATGGCCGGAACAGTGTTTCTAGATGGCAAAGAAGTAAACCAACTAAATGTCCAGTGG 3108

Qy    890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db   3109 CTCCGAGCACACCTAGGCATTGTGTCCCAAGAGCCCATCCTGTTTGACTGCAGCATCGCT 3168

Qy    910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db   3169 GAGAACATCGCTTATGGAGACAACAGCCGGGTCGTGTCCCAGGATGAGATCGAGAGGGCG 3228

Qy    930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db   3229 GCGAAGGAAGCCAACATCCACCAGTTCATTGAGTCACTGCCTGACAAATACAACACCAGA 3288

Qy    950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db   3289 GTAGGAGACAAAGGGACTCAGCTGTCGGGTGGGCAGAAACAACGCATCGCCATAGCACGC 3348

Qy    970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db   3349 GCCCTCGTCAGACAGCCTCACATTTTGCTCCTGGATGAAGCAACATCAGCCCTGGATACA 3408

Qy    990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db   3409 GAGAGTGAAAAGGTTGTCCAGGAAGCTCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 3468

Qy   1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db   3469 GTGATCGCTCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATTCAGAAT 3528

Qy   1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
Db   3529 GGCAAGGTCAAGGAGCATGGCACCCACCAGCAGCTGCTGGCACAGAAAGGCATCTATTTC 3588

Qy   1050 LysLeuValAsnAlaGlnSer 1056
Db   3589 TCCATGGTCAGTGTGCAGGCT 3609

RESULT 9
CRUPGP1165
LOCUS       CRUPGP1165            4304 bp    mRNA    linear   ROD 27-APR-1993
DEFINITION  Chinese hamster p-glycoprotein mRNA (clone ADX165), complete cds.
ACCESSION   M59253
VERSION     M59253.1  GI:191154
KEYWORDS    multidrug resistance glycoprotein; p-glycoprotein.
SOURCE      Chinese hamster, cDNA to mRNA, clone ADX165.
  ORGANISM  Cricetulus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
            Cricetulus.
REFERENCE   1  (bases 1 to 4304)
  AUTHORS   Devine,S.E., Hussain,A., Davide,J.P. and Melera,P.W.
  TITLE     Full length and alternatively spliced pgp1 transcripts in
            multidrug-resistant Chinese hamster lung cells
  JOURNAL   J. Biol. Chem. 266 (7), 4545-4555 (1991)
  MEDLINE   91154265
   PUBMED   1671863
FEATURES             Location/Qualifiers
     source          1..4304
                     /organism="Cricetulus sp."
                     /db_xref="taxon:10031"
                     /cell_type="fibroblast"
                     /tissue_type="lung"
     gene            1..4304
                     /gene="pgp-1"
     CDS             111..3941
                     /gene="pgp-1"
                     /codon_start=1
                     /product="p-glycoprotein"
                     /protein_id="AAA37004.1"
                     /db_xref="GI:191155"
                     /translation="MEFEEDFSGRKDKNFLKMGRKSKKEKKEKKPVVSVFTMFRYAGW
                     LDRLYMLVGTLAAIIHGVALPLMMLVFGDMTDSFASVGNIPTNATNNATQVNASDIFG
                     KLEEEMTTYAYYYTGIGAGVLIVAYIQVSFWCLAAGRQIHKIRQKFFHAIMNQEIGWF
                     DVHDVGELNTRLTDDVSKINEGIGDKIGMFFQAMATFFGGFIIGFTRGWKLTLVILAI
                     SPVLGLSAGIWAKILSSFTDKELQAYAKAGAVAEEVLAAIRTVIAFGGQKKELERYNN
                     NLEEAKRLGIKKAITANISMGAAFLLIYASYALAFWYGTSLVISKEYSIGQVLTVFFA
                     VLIAPFSIGQASPNIEAFANARGAAYEIFNIIDNKPSIDSFSKNGYKPDNIKGNLEFK
                     NIHFSYPSRKDVQILKGLNLKVQSGQTVALVGNSGCGKSTTVQLLQRLYDPTEGVVSI
                     DGQDIRTINVRYLREIIGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKEANAYD
                     FIMKLPHKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTESEAVV
                     QAALDKAREGRTTIVIAHRLSTVRNADIIAGFDGGVIVEQGNHEELMREKGIYFKLVM
                     TQTAGNEIELGNEVGESKNEIDNLDMSSKDSASSLIRRRSTRRSIRGPHDQDRKLSTK
                     EALDEDVPPISFWRILKLNSSEWPYFVVGIFCAIVNGALQPAFSIIFSKVVGVFTRNT
                     DDETKRHDSNLFSLLFLILGVISFITFFLQGFTFGKAGEILTKRLRYMVFKSMLRQDV
                     SWFDNPKNTTGALTTRLANDAGQVKGATGARLAVITQNIANLGTGIIISLIYGWQLTL
                     LLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGSGKIATEAIENFRTVVSLTREQKFE
                     NMYAQSLQIPYRNALKKAHVFGITFSFTQAMMYFSYAACFRFGAYLVARELMTFENVL
                     LVFSAIVFGAMAVGQVSSFAPDYAKAKVSASHIIMIIEKVPSIDSYSTGGLKPNTLEG
                     NVKFNEVVFNYPTRPDIPVLQGLNLEVKKGQTLALVGSSGCGKSTVVQLLERFYDPMA
                     GTVFLDGKEVNQLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSQDEIERAA
                     KEANIHQFIESLPDKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEATSALD
                     TESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVIQNGKVKEHGTHQQLLAQKG
                     IYFSMVSVQAGAKR"
BASE COUNT     1275 a    895 c   1031 g   1103 t
ORIGIN      Map position 1q26.

Alignment Scores:
Pred. No.:               3.04e-196     Length:          4304
Score:                   3136.50       Matches:         598
Percent Similarity:      75.82%        Conservative:    211
Best Local Similarity:   56.04%        Mismatches:      215
Query Match:             58.85%        Indels:          43
DB:                      10            Gaps:            6

US-09-873-409-4 (1-1058) x CRUPGP1165 (1-4304)

Qy      4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
Db    813 TCATTTACTGATAAAGAACTTCAGGCATATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTC 872

Qy     24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db    873 TTAGCAGCCATCAGAACTGTGATTGCATTTGGAGGACAAAAGAAAGAACTTGAA------ 926

Qy     44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db    926 ------------------------------------------------------------ 926

Qy     64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db    927 ---------------AGGTACAACAACAATTTAGAAGAAGCCAAAAGACTTGGGATAAAG 971

Qy     84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db    972 AAAGCGATCACAGCCAACATTTCCATGGGCGCAGCTTTCCTTCTTATCTATGCATCATAT 1031

Qy    104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db   1032 GCTCTGGCCTTCTGGTATGGAACGTCCTTGGTCATCTCAAAAGAA------TATTCTATT 1085

Qy    124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
```

```
Db   1086 GGACAAGTGCTCACTGTCTTCTTTGCTGTATTAATTGCGCCATTCAGTATTGGACAGGCA 1145

Qy    144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163

Db   1146 TCTCCAAATATCGAAGCCTTTGCCAATGCAAGAGGAGCAGCCTATGAAATCTTCAACATA 1205

Qy    164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183

Db   1206 ATTGATAATAAGCCCAGTATTGACAGCTTCTCAAAGAATGGGTACAAACCAGACAACATT 1265

Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203

Db   1266 AAAGGAAATTTGGAATTCAAAAATATTCACTTCAGTTACCCATCTCGAAAAGACGTTCAG 1325

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223

Db   1326 ATCTTGAAGGGCCTCAACCTGAAGGTGCAGAGCGGACAGACAGTGGCCCTGGTTGGCAAC 1385

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243

Db   1386 AGTGGCTGTGGGAAAAGCACCACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCACAGAG 1445

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263

Db   1446 GGCGTGGTCAGTATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAA 1505

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283

Db   1506 ATCATTGGGGTCGTGAGTCAGGAACCTGTGTTGTTTGCCACCACTATCGCTGAAAACATT 1565

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303

Db   1566 CGCTATGGCCGAGAAAATGTCACTATGGACGAGATCGAGAAAGCTGTCAAGGAAGCCAAT 1625

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323

Db   1626 GCCTATGACTTCATCATGAAACTGCCCCATAAATTTGACACTCTTGTTGGTGAGAGAGGA 1685

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343

Db   1686 GCACAGCTGAGTGGAGGACAGAAACAGAGAATCGCCATTGCTCGTGCCCTGGTCCGCAAC 1745

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363

Db   1746 CCCAAGATCCTTTTGTTGGATGAGGCGACATCAGCCTTGGACACAGAAAGTGAAGCCGTG 1805

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383

Db   1806 GTTCAGGCTGCTCTGGATAAGGCTAGAGAAGGCCGGACTACCATTGTGATAGCCCACCGC 1865

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403

Db   1866 TTGTCTACAGTTCGAAATGCTGACATCATAGCTGGGTTTGATGGTGGTGTCATTGTGGAG 1925

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423

Db   1926 CAAGGAAATCATGAAGAGCTCATGAGAGAGAAGGGCATTTACTTCAAACTTGTCATGACA 1985

Qy    424 Gln----------------------------AspIleLysLysAlaAspGluGlnMet 433

Db   1986 CAGACAGCAGGAAATGAAATTGAATTAGGAAATGAAGTTGGTGAGTCTAAAAATGAAATT 2045

Qy    434 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu------ProLeuHisSer 451

Db   2046 GATAATTTAGACATGTCTTCAAAAGATTCAGCATCCAGTCTAATCAGAAGGAGATCAACT 2105

Qy    452 ValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLys 470

Db   2106 CGCAGAAGTATCCGTGGACCACATGACCAAGACAGGAAGCTTAGTACCAAAGAGGCCTTG 2165

Qy    471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490

Db   2166 GATGAAGATGTACCTCCAATTTCCTTTTGGCGTATCCTAAAGTTGAATTCATCTGAATGG 2225

Qy    491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510

Db   2226 CCATATTTTGTGGTTGGTATATTCTGTGCCATAGTAAATGGAGCCTTGCAACCAGCATTC 2285

Qy    511 SerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThrThrLeu 529

Db   2286 TCAATAATATTTTCTAAGGTTGTAGGGGTTTTCACAAGAAATACTGATGATGAAACCAAA 2345

Qy    530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549

Db   2346 CGACATGATAGCAACTTGTTTTCCTTATTGTTTCTGATCCTTGGGGTCATTTCTTTTATT 2405

Qy    550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569

Db   2406 ACATTTTTTCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGACTC 2465

Qy    570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589

Db   2466 CGATACATGGTTTTCAAATCCATGCTGAGACAGGATGTGAGCTGGTTTGATAACCCTAAA 2525

Qy    590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609

Db   2526 AACACCACTGGAGCATTGACCACCAGGCTCGCCAACGATGCTGGTCAAGTTAAAGGGGCT 2585

Qy    610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629

Db   2586 ACAGGAGCCAGACTTGCTGTCATTACCCAGAATATAGCAAATCTTGGGACAGGAATCATC 2645

Qy    630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649

Db   2646 ATATCCCTAATCTATGGCTGGCAGTTGACACTTCTACTCTTAGCAATTGTCCCCATCATT 2705

Qy    650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669

Db   2706 GCAATAGCAGGAGTGGTTGAGATGAAAATGTTGTCTGGACAAGCACTAAAAGATAAGAAG 2765

Qy    670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689

Db   2766 GAGCTAGAAGGTTCTGGGAAGATTGCTACTGAAGCAATAGAGAACTTCCGCACTGTCGTC 2825

Qy    690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709

Db   2826 TCTTTGACTCGGGAGCAGAAGTTTGAAAATATGTATGCCCAGAGCTTGCAGATACCATAC 2885

Qy    710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729

Db   2886 AGAAATGCTCTGAAGAAAGCACACGTCTTTGGGATCACATTCTCCTTCACCCAGGCCATG 2945

Qy    730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749

Db   2946 ATGTATTTTTCGTATGCTGCTTGTTTCCGGTTTGGTGCTTACTTGGTGGCACGGGAATTA 3005

Qy    750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769

Db   3006 ATGACATTTGAAAATGTTCTATTAGTATTCTCAGCTATTGTCTTTGGTGCCATGGCAGTT 3065

Qy    770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789

Db   3066 GGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTGTCGGCATCGCATATC 3125

Qy    790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809

Db   3126 ATCATGATCATTGAAAAAGTCCCTTCCATTGACAGCTACAGCACCGGAGGCCTGAAGCCT 3185

Qy    810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829

Db   3186 AATACATTGGAAGGAAATGTGAAATTTAATGAAGTCGTGTTCAACTACCCCACCCGGCCT 3245

Qy    830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849

Db   3246 GACATCCCAGTGCTTCAGGGGCTGAACCTGGAGGTGAAGAAGGGCCAGACGCTGGCCCTG 3305
```

```
Qy      850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
            |||||||||||||||||||||||||||||||||   |||||||||||:::|||  ||||||
Db     3306 GTGGGCAGCAGTGGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAGCGGTTCTACGAC 3365

Qy      870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
            |||:::   |||   |||     ||||||   :::       :::||||||||||||||
Db     3366 CCCATGGCCGGAACAGTGTTTCTAGATGGCAAAGAAGTAAACCAACTAAATGTCCAGTGG 3425

Qy      890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
            ||||||:::   :::   ||||||   |||||||||:::||||||:::|||||||||||
Db     3426 CTCCGAGCACACCTAGGCATTGTGTCCCAAGAGCCCATCCTGTTTGACTGCAGCATCGCT 3485

Qy      910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
            |||||||||||||||||||||||||||||||||||||     |||||||||:::   |||
Db     3486 GAGAACATCGCTTATGGAGACAACAGCCGGGTCGTGTCCCAGGATGAGATCGAGAGGGCG 3545

Qy      930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
            |||      ||||||||||||   |||||||||   ||||||:::|||||||||||:::
Db     3546 GCGAAGGAAGCCAACATCCACCAGTTCATTGAGTCACTGCCTGACAAATACAACACCAGA 3605

Qy      950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
            ||||||   ||||||   ||||||||||||||||||||||||:::||||||||||||
Db     3606 GTAGGAGACAAAGGGACTCAGCTGTCGGGTGGGCAGAAACAACGCATCGCCATAGCACGC 3665

Qy      970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
            ||||||:::::::::|||   ||||||||||||||||||||||||||||||||||||
Db     3666 GCCCTCGTCAGACAGCCTCACATTTTGCTCCTGGATGAAGCAACATCAGCCCTGGATACA 3725

Qy      990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
            :::||||||||||||||||||   ||||||||||||||||||   ||||||||||||:::
Db     3726 GAGAGTGAAAAGGTTGTCCAGGAAGCTCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 3785

Qy     1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
            |||:::   ||||||||||||   ||||||||||||||||||||||||:::   |||
Db     3786 GTGATCGCTCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATTCAGAAT 3845

Qy     1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
            ||||||:::||||||   ||||||||||||:::||||||      :::   |||||||||
Db     3846 GGCAAGGTCAAGGAGCATGGCACCCACCAGCAGCTGCTGGCACAGAAAGGCATCTATTTC 3905

Qy     1050 LysLeuValAsnAlaGlnSer 1056
               :::|||:::   |||:::
Db     3906 TCCATGGTCAGTGTGCAGGCT 3926
```

```
RESULT 10
GGA9799
LOCUS       GGA9799                4018 bp    mRNA    linear   VRT 03-MAR-1999
DEFINITION  Gallus gallus mRNA for ABC transporter protein.
ACCESSION   AJ009799
VERSION     AJ009799.1  GI:3355756
KEYWORDS    ABC transporter protein; cmdr1 gene; P-glycoprotein.
SOURCE      Gallus gallus.
  ORGANISM  Gallus gallus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1  (bases 1 to 4018)
  AUTHORS   Edelmann,H.M.L.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-JUL-1998) Edelmann H.M.L., Institute of Molecular
            Genetics, Biocenter and University of Vienna, Dr. Bohr-Gasse 9/2
            A-1030 Vienna, A-1030 Vienna, AUSTRIA
REFERENCE   2  (bases 1 to 4018)
  AUTHORS   Edelmann,H.M., Duchek,P., Rosenthal,F.E., Foger,N., Glackin,C.,
            Kane,S.E. and Kuchler,K.
  TITLE     Cmdr1, a chicken P-glycoprotein, confers multidrug resistance and
            interacts with estradiol
  JOURNAL   Biol. Chem. 380 (2), 231-241 (1999)
  MEDLINE   99209805
   PUBMED   10195430
FEATURES             Location/Qualifiers
     source          1..4018
                     /organism="Gallus gallus"
                     /db_xref="taxon:9031"
                     /tissue_type="intestine"
     gene            109..3975
                     /gene="cmdr1"
     CDS             109..3975
                     /gene="cmdr1"
                     /note="ABC transporter protein"
                     /codon_start=1
                     /product="P-glycoprotein"
                     /protein_id="CAA08835.1"
                     /db_xref="GI:3355757"
                     /db_xref="SPTREMBL:O93437"
                     /translation="MHSEDEKQKHTVDGNTYEIANSNQDQDPEDEKKGKKKKGKKPQM
                     VSPLALFRYSSCTDKLLMIFGSLLAIAHGTSLPIAMIIFGDMTDSFVTSGMTNITGNS
                     SGLNSSADVFNKLEEEMTRYAYYYSAIAAAVLVAAYIQTSFWTLAAGRQVKKIREKFF
                     HAIMRQEIGWFDVNDAGELNTRLIDDVSKINEGIGDKIGFLIQSETTFLTGFIVGFIR
                     GWKLTLVILAVSPVLGLSAALWAKILTAFTDKEQAAYAKAGAVAEEVLSAVRTVIAFG
                     GQEKEIKRYHKNLEDAKRIGIRKAITSNISMGAAFLLIYASYALAFWYGTTLILANEY
                     SIGNVLTVFFSVLIGAFSIGQTAPSIEAFANARGAAYAIFNIIDNEPEIDSYSDAGHK
                     PDHIKGNLEFQNVFFNYPSRPDVEILKGLNLKVNCGQTVALVGGSGCGKSTTVQLIQR
                     FYDPKEGTITIDGQDLKSLNVRYLREIIGVVNQEPVLFATTIAENIRYGREDVTMEEI
                     ERATKEANAYDFIMKLPKKFETVVGERGAQMSGGQKQRIAIARALVHNPKILLLDEAT
                     SALDTESESVVQAALDKAREGRTTVVVAHRLSTVRNADLIAVFESGVITEQGNHSQLI
                     EKKGIYYKLVNMQTIETEDPSSEKSENAVSVKRSGSQSNLDESLKKELRRGSTRRSMK
                     KPGEPNDTDEKGSSPDEELPPVSFLKLMKLNKNEWPYFVAGTFCAIVNGALQPAFSVI
                     FSEIIGIFSETDQKVLREKSNLYSLLFLALGIISFFTFFVQGFAFGKAGEILTMKLRF
                     MAFKAMLRQDMAWFDDPKNSTGALTTRLANDASQVKGATGVRLALIAQNIANLGTGII
                     ISLVYGWQLTLLLLAVVPIIAVAGMIEMKMLAGHAKKDKIELEAAGKIATEAIENIRT
                     VASLTREKRFELMYGEHLLVPYRNSVKKAHIFGFCFSLSQAMMFFTYAGCFRFGAYLV
                     VNGHIEYKTVFLVFSAVVFGAMALGQTSSFAPDYAKAKISAAHLFVLFNRVPPIDSYR
                     EDGEKPEKFGGNTRIKDVKFNYPNRPEVKILQGLNLAVEKGETLALVGSSGCGKSTVV
                     QLLERFYDPLSGEIVFDDIDAKTLNIQWLRSHIGIVSQEPILFDFTIAENIAYGDNSR
                     EVSHEEIISAAKAASIHSFIDSLPEKYNTRVGDKGTQLSGGQKQRIAIARALIRKPQI
                     LLLDEATSALDTESEKIVQEALDKAREGRTCIVIAHRLSTIQNADKIAVIQNGKVIEQ
                     GTHQQLLAEKGFYYSLVNVQSGSCNM"
     mat_peptide     112..3972
                     /gene="cmdr1"
                     /product="ABC transporter protein"
                     /note="P-glycoprotein"
BASE COUNT     1288 a    758 c    911 g   1061 t
ORIGIN

Alignment Scores:
Pred. No.:              2.68e-195      Length:          4018
Score:                  3121.50        Matches:         599
Percent Similarity:     73.93%         Conservative:    192
Best Local Similarity:  55.98%         Mismatches:      236
Query Match:            58.56%         Indels:          43
DB:                     5              Gaps:            3

US-09-873-409-4 (1-1058) x GGA9799 (1-4018)

Qy        1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20
            ::::::   :::   |||   ||||||   :::||||||:::||||||||||||||||
Db      835 ATTCTCACTGCATTCACTGACAAAGAGCAAGCTGCATATGCAAAAGCAGGAGCTGTTGCA 894

Qy       21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40
            ||||||||||||||||:::::::|||||||||||||||||||      |||||||||||:::
Db      895 GAGGAAGTTCTGAGTGCAGTCCGTACTGTAATAGCTTTTGGAGGACAAGAAAAAGAAATT 954

Qy       41 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60
            :::
Db      955 AAA--------------------------------------------------------- 957

Qy       61 LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80
                                    ||||||   :::||||||:::||||||||
Db      958 -----------------------AGATATCATAAAAATTTAGAAGATGCTAAGCGCATT 993

Qy       81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 100
            ||||||::::::   |||   |||   :::|||:::||||||      :::      :::
Db      994 GGAATAAGAAAGGCTATTACATCTAATATTTCTATGGGTGCTGCTTTCTTACTGATATAT 1053
```

```
Qy   101 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120
Db  1054 GCATCATATGCACTGGCCTTCTGGTATGGAACTACCTTGATCTTAGCCAATGAG------ 1107

Qy   121 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 140
Db  1108 TACAGTATTGGCAATGTTCTTACAGTCTTTTTCTCTGTACTGATTGGAGCTTTCAGTATT 1167

Qy   141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
Db  1168 GGACAGACTGCTCCAAGCATAGAGGCATTTGCTAATGCAAGAGGAGCAGCCTATGCAATC 1227

Qy   161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180
Db  1228 TTTAATATCATAGACAATGAACCTGAAATTGACAGTTATTCAGATGCAGGCCACAAACCA 1287

Qy   181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200
Db  1288 GACCACATTAAAGGGAATTTGGAATTCCAAAATGTTTTCTTCAATTACCCATCCAGACCA 1347

Qy   201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
Db  1348 GATGTTGAGATACTGAAAGGCTTGAATCTCAAGGTTAATTGCGGCCAGACAGTAGCCCTG 1407

Qy   221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240
Db  1408 GTTGGTGGCAGCGGCTGTGGGAAAAGTACAACTGTTCAGCTCATCCAGAGATTTTACGAC 1467

Qy   241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260
Db  1468 CCCAAGGAAGGCACGATTACCATTGACGGGCAGGATCTGAAGAGCCTGAATGTAAGATAT 1527

Qy   261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280
Db  1528 CTGCGAGAGATTATTGGTGTGGTGAACCAGGAGCCCGTGCTGTTTGCTACAACTATTGCA 1587

Qy   281 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 300
Db  1588 GAAAATATTCGTTATGGCCGTGAGGATGTCACCATGGAAGAAATTGAAAGAGCTACCAAG 1647

Qy   301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320
Db  1648 GAAGCTAATGCTTATGATTTCATCATGAAGCTCCCCAAGAAGTTTGAAACTGTGGTTGGA 1707

Qy   321 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340
Db  1708 GAAAGAGGGGCACAGATGAGTGGAGGGCAGAAGCAGCGAATAGCAATTGCCCGTGCTCTG 1767

Qy   341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360
Db  1768 GTTCACAATCCTAAAATTCTTCTGCTTGATGAGGCAACGTCAGCTTTGGATACTGAGAGT 1827

Qy   361 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380
Db  1828 GAATCCGTTGTGCAGGCAGCACTGGACAAGGCAAGAGAAGGACGGACCACGGTTGTAGTA 1887

Qy   381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400
Db  1888 GCTCATCGTCTATCAACAGTCCGAAATGCAGATCTTATCGCTGTGTTTGAAAGTGGAGTT 1947

Qy   401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420
Db  1948 ATCACAGAACAAGGAAATCATTCTCAACTGATTGAAAAAAAGGGAATCTACTATAAACTT 2007

Qy   421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440
Db  2008 GTTAACATGCAGACAATAGAAACTGAAGATCCATCATCAGAAAAAAGTGAAAATGCAGTC 2067

Qy   441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPhe--- 459
Db  2068 AGTGTCAAAAGAAGTGGATCTCAATCAAATCTGGATGAATCCCTAAAAAAAGAATTGAGA 2127

Qy   460 ---------------------------------------IleAspLysAlaGluGluSer 466
Db  2128 AGAGGATCAACTAGGAGAAGCATGAAAAAGCCAGGAGAACCTAATGATACTGATGAGAAA 2187

Qy   467 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 486
Db  2188 GGTAGTTCACCTGATGAAGAATTACCTCCAGTTTCATTTCTGAAACTTATGAAGTTGAAC 2247

Qy   487 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 506
Db  2248 AAAAACGAATGGCCATACTTTGTAGCTGGAACTTTCTGTGCAATTGTCAATGGAGCTTTA 2307

Qy   507 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys 526
Db  2308 CAACCTGCATTTTCAGTCATATTCTCAGAAATTATAGGGATTTTTCAGAAACTGATCAG 2367

Qy   527 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 546
Db  2368 AAAGTTTTGAGAGAAAAGAGCAACTTATATTCACTGCTGTTTTAGCACTTGGGATCATT 2427

Qy   547 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 566
Db  2428 TCCTTTTTACTTCTTTGTTCAGGGTTTTGCATTTGGCAAAGCTGGAGAAATTCTTACA 2487

Qy   567 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 586
Db  2488 ATGAAGCTTCGATTCATGGCATTTAAAGCAATGCTGAGACAGGACATGGCCTGGTTTGAT 2547

Qy   587 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 606
Db  2548 GATCCTAAAAATAGCACTGGAGCACTGACAACAAGACTTGCTAATGATGCCTCACAAGTA 2607

Qy   607 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 626
Db  2608 AAAGGAGCAACTGGTGTCAGATTGGCATTAATTGCCCAAAACATAGCTAACCTGGGAACT 2667

Qy   627 SerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAla 646
Db  2668 GGAATTATTATATCTTTAGTTTATGGCTGGCAGCTGACCCTGTTACTCTTAGCTGTTGTA 2727

Qy   647 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 666
Db  2728 CCAATCATTGCAGTGGCAGGAATGATTGAAATGAAGATGTTGGCTGGGCATGCAAAAAAA 2787

Qy   667 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 686
Db  2788 GATAAAATAGAGCTGGAAGCTGCAGGAAAGATTGCCACAGAAGCCATAGAAAATATTCGA 2847

Qy   687 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 706
Db  2848 ACTGTTGCTAGTTTGACTCGGGAAAAAAGATTTGAGCTCATGTACGGAGAACATTTGCTT 2907

Qy   707 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 726
Db  2908 GTACCATACAGAAATTCTGTGAAGAAGGCGCACATATTTGGGTTTGTTTTTCCCTTTCA 2967

Qy   727 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 746
Db  2968 CAAGCAATGATGTTTTTTACCTATGCTGGCTGTTTCCGCTTTGGAGCCTATCTTGTGGTT 3027

Qy   747 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 766
Db  3028 AATGGACACATAGAGTACAAAACTGTGTTTTTAGTGTTTTCAGCTGTTGTATTTGGTGCT 3087

Qy   767 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 786
Db  3088 ATGGCATTAGGACAAACCAGCTCTTTTGCACCAGACTATGCCAAAGCCAAGATATCAGCA 3147

Qy   787 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 806
Db  3148 GCCCACTTGTTTGTGCTGTTTAATAGAGTACCTCCAATAGACAGTTACCGCGAGGATGGA 3207

Qy   807 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 826
```

```
Db   3208 GAGAAACCTGAAAAATTTGGGGGAAACACAAGGATCAAAGATGTGAAATTTAACTATCCA 3267

Qy    827 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 846
             ||||||:::|||  ||||||:::||||||:::|||::::::|||:::|||:::|||
Db   3268 AACCGACCAGAGGTCAAAATCCTCCAAGGCTTGAATCTAGCAGTAGAAAAAGGAGAAACG 3327

Qy    847 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 866
          :::|||   |||||||||||||||||||||||||||||   ||||||||||||:::|||
Db   3328 TTGGCCCTTGTTGGTAGCAGTGGATGTGGAAAGAGCACTGTTGTTCAGCTACTTGAGAGA 3387

Qy    867 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 886
             |||||||||:::   |||:::::::::||||||   :::||||||||||   ||||||
Db   3388 TTTTATGATCCACTCAGTGGAGAAATTGTGTTTGATGACATTGATGCAAAGACACTAAAT 3447

Qy    887 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 906
          :::||||||||||||||||   |||   ||||||   ||||||||:::||||||:::
Db   3448 ATCCAGTGGCTGAGATCTCACATCGGTATCGTCTCACAAGAGCCAATCCTGTTCGACTTC 3507

Qy    907 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 926
          :::|||||||||||||||||||||||||||||||||||   |||      :::||||||
Db   3508 ACCATCGCTGAAAACATCGCGTACGGGGACAACAGTCGGGAGGTGTCCCATGAGGAAATC 3567

Qy    927 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 946
             ||||||   ||||||:::|||||||||||||||:::   ||||||||||||||||
Db   3568 ATTAGTGCAGCGAAAGCAGCCAGTATTCATTCCTTCATTGACTCTCTGCCAGAAAAATAC 3627

Qy    947 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 966
          ||||||:::|||||||||   |||||||||||||||||||||||||||||||:::|||
Db   3628 AATACTCGTGTAGGAGACAAGGGAACACAGCTTTCTGGTGGTCAGAAACAACGTATTGCT 3687

Qy    967 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 986
          |||||||||||||||::::::||||||:::|||||||||||||||||||||||||||||
Db   3688 ATTGCTCGAGCTCTTATACGTAAGCCCCAGATTCTGCTACTGGATGAAGCTACATCTGCC 3747

Qy    987 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 1006
          ||||||   :::|||||||||:::||||||   ||||||||||||||||   ||||||
Db   3748 TTGGACACAGAAAGCGAAAAGATTGTCCAGGAAGCACTGGATAAAGCCAGAGAAGGTCGC 3807

Qy   1007 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 1026
          ||||||:::|||:::   ||||||||||||   ||||||||||||||||   |||   |||
Db   3808 ACCTGCATCGTGATAGCTCACCGCCTGTCCACCATCCAGAATGCTGACAAGATTGCTGTG 3867

Qy   1027 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 1046
          :::   |||||||||:::   ||||||||||||||||||||:::||||||      :::
Db   3868 ATCCAGAACGGCAAGGTCATAGAGCAAGGGACTCACCAGCAGCTCCTGGCTGAAAAAGGC 3927

Qy   1047 IleTyrPheLysLeuValAsnAlaGlnSer 1056
             |||:::   |||||||||   ||||||
Db   3928 TTCTACTACTCGCTGGTCAATGTTCAAAGT 3957
```

```
RESULT 11
OAU78609
LOCUS       OAU78609                3858 bp    mRNA    linear   MAM 11-OCT-2001
DEFINITION  Ovis aries multidrug resistance protein-1 (mdr1) mRNA, complete
            cds.
ACCESSION   U78609
VERSION     U78609.1  GI:2149086
KEYWORDS    .
SOURCE      Ovis aries.
  ORGANISM  Ovis aries
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
            Bovidae; Caprinae; Ovis.
REFERENCE   1  (bases 1 to 3858)
  AUTHORS   Longley,M., Phua,S.H., van Stijn,T.C. and Crawford,A.M.
  TITLE     Isolation and mapping of the first ruminant multidrug resistance
            genes
  JOURNAL   Anim. Genet. 30 (3), 207-210 (1999)
  MEDLINE   99371931
   PUBMED   10442984
REFERENCE   2  (bases 1 to 3858)
  AUTHORS   Longley,M. and Crawford,A.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-NOV-1996) Biochemistry Department, University of
            Otago, Dunedin, New Zealand
FEATURES             Location/Qualifiers
     source          1..3858
                     /organism="Ovis aries"
                     /db_xref="taxon:9940"
     gene            1..3858
                     /gene="mdr1"
     CDS             1..3858
                     /gene="mdr1"
                     /codon_start=1
                     /product="multidrug resistance protein-1"
                     /protein_id="AAB58489.1"
                     /db_xref="GI:2149087"
                     /translation="MDLEGDRNGRAAGGNFLKRDKKRFFSKKDEKKEKRPTVSTFTMF
                     RYSNWLDRLCMVLGTLAAIIHGAGLPLMMLVFGDMTDSFAGAGNLGNITLSNISNTST
                     IDRTEYGKKLEKEMTTYAYYYSGIGAGVLIAAYIQVSFWCLAAGRQVHRIRKQFFHAI
                     MQQEIGWFDVHDVGKLNTRLTNDVSKINEGIGDKIGMFFQAMATFFTGFIIGFTTGWN
                     LTLVILAIIPVLGLSAAIWAKILSSFTDKKLLAYAKAGAVAEEVLAAIKTVIAFGGQK
                     KELERYNKNLEEAKRIGIKKAITANISMGAAFLLIYASYALAFWYGTSLVLSREYSIG
                     QVLTVFFSVLIGAFSIGQASPNIEAFANARGAAYEVFKIIDNKPSIDSYSNTGHKPDN
                     IKGNLEFRNVHFHYPSRNEVKILKGLNLKVGSGQTVALVGNSGCGKSTTVQLMQRLYD
                     PTEGMVSIDGQDIRTINVRYLREIIGVVSQEPVLFATTIAENIRYGREDVTMDEIQKA
                     VKEANAYDFIMKLPNKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSAL
                     GTESEAVVQAALDKARKGRTTIVIAHRLSTVRNADIIAGLDDGVIVEEGSHDELMGKR
                     GIYFKLVTMQTKGNELELENTPGESLSKTDDLYMSSQDSRSSLIRRKSTRRSIRGSQS
                     QDRKLSTEETLDESVPPVSFWRILKLNITEWPYFVVGVFCAIINGALQPAFSVIFSRI
                     IGIFTRNDNDETKRQNSNLFSLLFLILGIISFITFFLQGFTFGKAGEILTRRLRYLVF
                     RSMLRQDVSWFDDPKNTTGALTTRLANDAAQVKGAVGSRLAVITQNIANLGTGIIISL
                     IYGWQLTLLLLAIVPIIAVAGVIEMKMLSGQALKDKKELEGAGKIATEAIENFRTVVS
                     LTREERFEYMYAQSLQVPYRNSLRRAHVFGITFSITQAMMYFSYAGCFRFGAYLVAQG
                     IMEFQDVLLVFSAVVFGAMAVGQVSTFAPDYAKAKVSAAHVINIIEKIPLIDSYSTEG
                     LKPSTVEGSVAFNDVVFNYPTRPDVPVLRGLSLEVKKGQTLALVGSSGCGKSTVVQLL
                     ERFYDPLAGTVFIDGKEVKQLNVQWLRAHMGIVSQEPILFDCSIGENIAYGDNSRVVS
                     QEEIEHAAKEANIHSFIEMLPDKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLL
                     DEATSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGRIKEHGTH
                     QQLLAQKGIYFTMVSVQAGTKRQ"
BASE COUNT     1072 a    853 c   1002 g    931 t
ORIGIN

Alignment Scores:
Pred. No.:              4.01e-195     Length:         3858
Score:                  3118.50       Matches:        602
Percent Similarity:     75.30%        Conservative:   203
Best Local Similarity:  56.31%        Mismatches:     217
Query Match:            58.51%        Indels:         47
DB:                     4             Gaps:           6

US-09-873-409-4 (1-1058) x OAU78609 (1-3858)

Qy      4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
          |||   |||   |||:::|||   ||||||:::||||||||||||||||||||||||
Db    727 TCATTTACTGATAAAAAACTTTTGGCATATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 786

Qy     24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
          |||::::::|||:::|||||||||||||||      |||:::|||||||||:::
Db    787 TTAGCGGCTATTAAAACTGTGATTGCATTTGGAGGACAAAAAAAAGAACTTGAA------ 840

Qy     44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63

Db    840 ------------------------------------------------------------ 840

Qy     64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
                         ||||||   :::||||||::::::||||||      |||||||||
Db    841 ---------------AGGTACAACAAAAACTTGGAAGAAGCCAAGAGAATTGGAATAAAG 885

Qy     84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
          :::   |||   :::   :::|||:::||||||   :::      :::      :::|||
Db    886 AAGGCTATCACGGCCAACATCTCTATGGGAGCTGCTTTTCTGTTGATCTATGCATCCTAT 945

Qy    104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
             ||||||||||||||||||||||||:::|||:::   |||      |||:::|||
Db    946 GCCCTGGCGTTCTGGTACGGGACCTCCTTGGTCCTCTCAAGAGAA------TATTCCATT 999
```

```
Qy     124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db    1000 GGACAAGTGCTCACAGTCTTCTTTTCTGTATTAATTGGAGCTTTAGTATTGGACAGGCA 1059

Qy     144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db    1060 TCTCCAAACATTGAGGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAGTCTTCAAGATC 1119

Qy     164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db    1120 ATTGATAACAAACCAAGCATCGACAGCTATTCAAACACTGGGCACAAACCCGACAATATT 1179

Qy     184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db    1180 AAAGGAAATCTGGAATTCAGAAATGTTCACTTCCATTACCCATCTCGAAACGAAGTTAAG 1239

Qy     204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db    1240 ATCTTGAAGGGCCTCAACCTGAAGGTAGGGAGCGGGCAGACGGTGGCCCTGGTTGGGAAC 1299

Qy     224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db    1300 AGTGGCTGTGGGAAGAGCACCACCGTCCAGCTGATGCAGAGGCTGTATGACCCCACGGAG 1359

Qy     244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db    1360 GGCATGGTCAGTATCGACGGACAGGATATCAGGACCATCAACGTAAGGTATCTGCGGGAG 1419

Qy     264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db    1420 ATTATCGGGGTGGTGAGTCAGGAGCCTGTGCTGTTTGCCACCACGATAGCTGAGAACATT 1479

Qy     284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db    1480 CGCTACGGCCGTGAAGATGTCACCATGGATGAGATTCAGAAAGCTGTGAAGGAAGCCAAC 1539

Qy     304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db    1540 GCCTATGACTTTATCATGAAACTGCCTAATAAGTTTGACACCCTGGTTGGAGAGAGAGGA 1599

Qy     324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db    1600 GCCCAGCTAAGCGGTGGACAGAAGCAAAGAATCGCCATTGCCCGGGCCCTGGTTCGCAAC 1659

Qy     344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db    1660 CCCAAGATCCTCCTGTTGGATGAGGCCACCTCAGCGCTGGGCACCGAGAGTGAAGCGGTG 1719

Qy     364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db    1720 GTTCAGGCGGCCCTGGATAAGGCCAGAAAAGGCCGGACCACCATTGTGATAGCGCATCGT 1779

Qy     384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db    1780 CTGTCCACGGTTCGCAATGCTGACATCATTGCTGGTCTTGATGATGGAGTCATCGTGGAG 1839

Qy     404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db    1840 GAAGGCAGTCATGATGAGCTCATGGGGAAAAGAGGCATTTACTTCAAACTTGTCACAATG 1899

Qy     424 GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 443
Db    1900 CAG---ACAAAAGGAAATGAACTTGAATTGGAAAATACCCCTGGTGAATCCCTAAGTAAA 1956

Qy     444 ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--------- 460
Db    1957 ACTGATGACTTGTACATGTCCTCCCAAGATTCA---AGATCCAGTCTAATCAGAAGAAAA 2013

Qy     461 -----------------------------------AspLysAlaGluGluSerThrGln 468
Db    2014 TCAACTCGCAGGAGTATCCGTGGATCCCAAAGCCAGGACAGAAAGCTCAGTACAGAAGAA 2073

Qy     469 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 488
Db    2074 ACTTTGGATGAAAGTGTACCTCCCGTTTCCTTTTGGAGGATTCTGAAGCTGAATATAACT 2133

Qy     489 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 508
Db    2134 GAATGGCCTTATTTGTGGTTGGTGTATTTGCGCCATTATAAATGGAGCTCTGCAACCA 2193

Qy     509 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLys---Thr 527
Db    2194 GCATTCTCAGTCATATTTTCAAGGATTATAGGAATCTTCACCAGAAATGACAATGACGAA 2253

Qy     528 ThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 547
Db    2254 ACCAAACGACAGAACAGTAACTTGTTTTCACTGTTGTTTCTAATCCTTGGAATTATTTCT 2313

Qy     548 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 567
Db    2314 TTTATTACATTTTTCCTTCAGGGCTTCACGTTCGGCAAAGCAGGCGAGATCCTCACCAGG 2373

Qy     568 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 587
Db    2374 CGGCTGCGGTACCTGGTTTTCAGGTCGATGCTGAGACAGGATGTGAGCTGGTTCGATGAC 2433

Qy     588 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 607
Db    2434 CCTAAAAACACCACGGGAGCATTGACGACCAGGCTTGCCAATGACGCCGCTCAAGTTAAA 2493

Qy     608 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 627
Db    2494 GGGGCTGTAGGTTCCAGGCTTGCTGTAATTACCCAGAATATTGCAAATCTTGGCACAGGC 2553

Qy     628 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 647
Db    2554 ATTATCATATCCCTCATCTATGGCTGGCAGTTAACACTTCTCCTGTTAGCAATTGTACCC 2613

Qy     648 ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp 667
Db    2614 ATCATTGCAGTAGCAGGAGTTATTGAAATGAAAATGCTGTCTGGACAAGCCCTGAAAGAT 2673

Qy     668 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 687
Db    2674 AAGAAAGAACTGGAAGGTGCTGGGAAGATTGCCACTGAAGCAATAGAAAACTTCCGGACG 2733

Qy     688 IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 707
Db    2734 GTTGTTTCTTTGACTCGGGAGGAGAGGTTTGAATATATGTATGCCCAGAGTTTGCAGGTA 2793

Qy     708 GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis 727
Db    2794 CCATACAGAAACTCCTTGAGGAGAGCGCATGTCTTTGGAATTACATTTTCCATCACTCAG 2853

Qy     728 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla 747
Db    2854 GCAATGATGTATTTCTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTGGCCCAA 2913

Qy     748 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 767
Db    2914 GGCATCATGGAGTTTCAGGATGTTCTCTTAGTGTTCTCAGCTGTTGTCTTTGGTGCCATG 2973

Qy     768 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 787
Db    2974 GCTGTGGGACAGGTCAGTACATTTGCTCCTGACTATGCCAAGGCCAAAGTGTCAGCAGCC 3033

Qy     788 HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys 807
Db    3034 CACGTCATCAATATCATTGAAAAAATTCCTCTGATTGACAGCTACAGCACGGAAGGCCTA 3093

Qy     808 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys 827
Db    3094 AAGCCGAGTACAGTGGAAGGAAGTGTGGCATTTAATGACGTCGTGTTCAACTACCCCACT 3153

Qy     828 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 847
```

```
               |||||||||||   :::||||||||||||||||   ::::::::|||:::|||:::
Db       3154 CGACCAGACGTCCCAGTGCTTCGGGGGCTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTG 3213

Qy        848 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 867
              |||  ||||||||||||||||||||||||||||||  |||||||||||:::|||
Db       3214 GCGCTGGTGGGCAGTAGTGGCTGTGGGAAGAGCACGGTCGTCCAGCTCCTGGAGCGGTTC 3273

Qy        868 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 887
              |||||||||:::  |||  |||      |||||   :::    |||:::||||||||
Db       3274 TATGACCCCTTGGCTGGCACAGTGTTTATTGACGGCAAGGAAGTAAAGCAGCTGAATGTC 3333

Qy        888 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 907
              |||||||||||:::  :::  |||||  ||||||||:::||||||:::||||||
Db       3334 CAGTGGCTGCGGGCACACATGGGCATCGTGTCCCAGGAGCCCATTCTGTTTGACTGCAGC 3393

Qy        908 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 927
              |||  ||||||||||||||||||||||||||||||||||||      :::||||||:::
Db       3394 ATCGGCGAGAATATCGCCTACGGAGACAACAGCCGGGTCGTGTCCCAGGAGGAGATTGAG 3453

Qy        928 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn 947
                    ||||||     ||||||||||||||||||||||||   ||||||:::||||||||
Db       3454 CACGCGGCCAAGGAGGCCAACATCCACTCCTTCATCGAGATGCTGCCGGACAAATACAAC 3513

Qy        948 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 967
              |||:::||||||  ||||||||||||||||||||||||||||||||||:::||||||
Db       3514 ACCAGAGTAGGGGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAGCGCATTGCCATA 3573

Qy        968 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 987
              ||||||||||||:::::::::|||  ||||||||||||||||||||||||||||||||
Db       3574 GCTCGCGCCCTTGTTAGACAGCCTCACATCTTGCTTCTGGATGAAGCTACCTCGGCTCTG 3633

Qy        988 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1007
              |||   :::|||||||||||||||||||  ||||||||||||||||||  ||||||||
Db       3634 GACACGGAAAGTGAAAAGGTGGTCCAAGAAGCCCTGGACAAAGCCCGAGAGGGACGCACC 3693

Qy       1008 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1027
              |||:::|||:::   |||||||||||   |||||||||||||||||||||||||||
Db       3694 TGCATCGTGATCGCTCACCGGCTGTCCACCATCCAGAACGCAGACTTGATAGTGGTGTTT 3753

Qy       1028 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1047
                 ||||||:::||||||||  ||||||||||||:::||||||       :::   |||
Db       3754 CAGAATGGCAGAATCAAGGAGCATGGCACACACCAGCAGCTGCTGGCGCAGAAAGGCATC 3813

Qy       1048 TyrPheLysLeuValAsnAlaGlnSer 1056
              ||||||   :::|||:::   |||:::
Db       3814 TATTTCACCATGGTCAGTGTCCAGGCT 3840
```

```
RESULT 12
AF257746
LOCUS       AF257746                4927 bp    mRNA    linear   ROD 09-MAY-2000
DEFINITION  Rattus norvegicus multidrug resistance protein 1a (Pgy1) mRNA,
            complete cds.
ACCESSION   AF257746
VERSION     AF257746.1  GI:7739772
KEYWORDS    .
SOURCE      Rattus norvegicus.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 4927)
  AUTHORS   Hooiveld,G.J.E.J., Wilms,J.W.J., Hagenbuch,B., Jansen,P.L.M.,
            Meijer,D.K.F. and Muller,M.
  TITLE     Cloning and functional characterization of the rat multidrug
            resistance protein Mdr1a
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4927)
  AUTHORS   Hooiveld,G.J.E.J., Wilms,J.W.J., Hagenbuch,B., Jansen,P.L.M.,
            Meijer,D.K.F. and Muller,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-APR-2000) Div. Gastroenterology & Hepatology,
            University Hospital Groningen, Hanzeplein 1, Groningen NL-9713 GZ,
            The Netherlands
FEATURES             Location/Qualifiers
     source          1..4927
                     /organism="Rattus norvegicus"
                     /strain="Wistar"
                     /db_xref="taxon:10116"
                     /sex="male"
                     /cell_type="hepatocyte"
                     /tissue_type="liver"
                     /dev_stage="adult"
     gene            1..4927
                     /gene="Pgy1"
     CDS             86..3904
                     /gene="Pgy1"
                     /note="ATP-binding cassette transport protein; Abcb1a"
                     /codon_start=1
                     /product="multidrug resistance protein 1a"
                     /protein_id="AAF69007.1"
                     /db_xref="GI:7739773"
                     /translation="MELEEDLNGRADKNFSKMGKKSKKEKKEKKPAVSVLTMFRYAGW
                     LDRFYMLLGTLAAIIHGIALPLMMLVFGDMTDSFANVGNNRSMSFYNATDIYAKLEDE
                     MTTYAYYYTGIGAGVLIVAYIQVSLWCLAAGRQIHKIRQKFFHAIMNQEIGWFDVHDV
                     GELNTRLTDDVSKINEGIGDKIGMFFQAMATFFGGFIIGFTRGWKLTLVILAISPVLG
                     LSAGIWAKILSSFTDKELQAYAKAGAVAEEVLAAIRTVIAFGGQKKELERYNNNLEEA
                     KRLGIKKAITANISMGAAFLLIYASYALAFWYGTSLVISKEYTIGQVLTVFFSVLIGA
                     FSVGQASPNIEAFANARGAAYEVFSIIDNKPSIDSFSKSGHKPDNIQGNLEFKNIHFS
                     YPSRKDVQILKGLNLKVKSGQTVALVGNSGCGKSTTVQLLQRLYDPIEGEVSIDGQDI
                     RTINVRYLREIIGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKEANAYDFIMKL
                     PHKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTESEAVVQAALD
                     KAREGRTTIVIAHRLSTVRNADVIAGFDGGVIVEQGNHDELMREKGIYFKLVMTQTAG
                     NEIELGNEACESKDGIDNVDMSSKDSGSSLIRRRSTRKSIRGPHDQDGELSTKEALDD
                     DVPPASFWRILKLNSTEWPYFVVGVFCAIINGGLQPAFSIIFSKVVGVFTKNDTPEIQ
                     RQNSNLFSLLFLILGIISFITFFLQGFTFGKAGEILTKRLRYMVFKSMLRQDISWFDD
                     PKNTTGALTTRLANDAAQVKGATGSRLAVITQNIANLGTGIIISLIYGWQLTLLLLAI
                     VPIIAIAGVVEMKMLSGQALKDKKELEGSGKIATEAIENFRTVVSLTREQKFETMYAQ
                     SLQIPYRNALKKAHVFGITFSFTQAMMYFSYAACFRFGAYLVARELMTFENVLLVFSA
                     IVFGAMAVGQVSSFAPDYAKAKVSASHIIRIIEKIPEIDSYSTEGLKPNMLEGNVKFN
                     GVMFNYPTRPNIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERFYDPMAGTVFL
                     DGKEIKQLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSHEEIVKAAKEANI
                     HQFIDSLPEKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEATSALDTESEK
                     VVQEALDKAREGRTCIVIAHRLSTIQNADLIVVIQNGQVKEHGTHQQLLAQKGIYFSM
                     VSVQAGAKRS"
BASE COUNT     1474 a   1037 c   1190 g   1226 t
ORIGIN

Alignment Scores:
Pred. No.:                7.34e-195       Length:          4927
Score:                    3116.50         Matches:         601
Percent Similarity:       75.73%          Conservative:    207
Best Local Similarity:    56.33%          Mismatches:      216
Query Match:              58.47%          Indels:          43
DB:                       10              Gaps:            7

US-09-873-409-4 (1-1058) x AF257746 (1-4927)

Qy          4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
              |||   |||   ||||||||   ||||||:::||||||||||||||||||||||||
Db        773 TCATTTACTGATAAGGAACTCCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTC 832

Qy         24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
              |||::::::||||||||||||||||||         |||:::||||||||:::
Db        833 TTAGCAGCCATCAGAACTGTGATTGCCTTTGGAGGACAAAAGAAGGAACTTGAA------ 886

Qy         44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63

Db        886 ------------------------------------------------------------ 886

Qy         64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
                           ||||||      ||||||::::::||||||      ||||||||
Db        887 ---------------AGGTACAATAACAATTTGGAAGAAGCTAAAAGGCTTGGGATAAAG 931

Qy         84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
              :::   |||   :::   :::|||:::||||||   :::       :::       :::|||
Db        932 AAAGCTATCACGGCCAACATTTCCATGGGTGCAGCTTTTCTGCTTATCTATGCATCATAT 991
```

```
Qy     104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db     992 GCTCTGGCATTCTGGTATGGGACTTCCTTGGTCATCTCAAAAGAA------TACACTATT 1045

Qy     124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db    1046 GGACAAGTGCTCACTGTCTTTTTTTCTGTATTAATTGGAGCATTCAGTGTTGGGCAGGCA 1105

Qy     144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db    1106 TCTCCAAATATTGAAGCCTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTCAGTATA 1165

Qy     164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db    1166 ATTGATAATAAGCCCAGTATAGACAGCTTCTCAAAGAGTGGGCACAAACCCGACAACATA 1225

Qy     184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db    1226 CAAGGAAATTTGGAATTCAAAAATATTCACTTCAGTTACCCGTCTCGAAAAGACGTTCAG 1285

Qy     204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db    1286 ATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTGGTTGGCAAC 1345

Qy     224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db    1346 AGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCATAGAG 1405

Qy     244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db    1406 GGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAA 1465

Qy     264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db    1466 ATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCCACCACAATTGCCGAAAACATT 1525

Qy     284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db    1526 CGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAAT 1585

Qy     304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db    1586 GCCTATGATTTCATCATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAGAGAGGG 1645

Qy     324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db    1646 GCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAAC 1705

Qy     344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db    1706 CCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTG 1765

Qy     364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db    1766 GTTCAGGCCGCTCTGGATAAGGCTAGAGAAGGCCGGACCACCATTGTGATAGCTCACCGC 1825

Qy     384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db    1826 TTGTCTACAGTTCGCAATGCTGACGTCATTGCTGGTTTTGATGGTGGTGTCATTGTGGAG 1885

Qy     404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db    1886 CAAGGAAATCATGATGAGCTCATGAGAGAGAAAGGAATTTACTTCAAACTTGTCATGACT 1945

Qy     424 Gln------------AspIleLysLysAlaAspGluGlnMetGluSer------------ 435
Db    1946 CAGACAGCAGGAAATGAAATTGAATTAGGAAATGAAGCTTGTGAATCTAAAGATGGAATT 2005

Qy     436 ------MetThrTyrSerThrGluArgLysThrAsnSerLeu------ProLeuHisSer 451
Db    2006 GATAATGTGGACATGTCTTCAAAAGATTCAGGATCCAGTCTAATAAGAAGAAGATCAACT 2065

Qy     452 ValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLys 470
Db    2066 CGCAAAAGCATCCGTGGGCCACATGATCAAGACGGGGAACTTAGCACCAAAGAGGCTCTG 2125

Qy     471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
Db    2126 GATGACGACGTACCTCCAGCTTCCTTTTGGCGGATCCTGAAGTTGAATTCAACTGAATGG 2185

Qy     491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db    2186 CCTTATTTTGTGGTTGGTGTATTTTGTGCCATAATAAATGGAGGCTTGCAACCAGCATTC 2245

Qy     511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeu--- 529
Db    2246 TCCATAATATTTTCAAAGGTTGTAGGGGTTTTTACAAAAAATGACACCCCTGAAATCCAG 2305

Qy     530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db    2306 CGGCAGAACAGCAACTTGTTTTCTTTATTGTTCCTGATCCTTGGGATCATCTCTTTCATT 2365

Qy     550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db    2366 ACGTTTTTCCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGACTC 2425

Qy     570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
Db    2426 CGATACATGGTCTTCAAATCCATGCTGAGACAGGACATAAGCTGGTTTGATGACCCTAAA 2485

Qy     590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db    2486 AACACCACAGGAGCGCTGACCACCAGGCTTGCCAATGACGCTGCTCAAGTGAAAGGGGCT 2545

Qy     610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db    2546 ACAGGGTCTAGGCTTGCTGTTATTACCCAGAACATAGCAAATCTTGGGACAGGCATCATC 2605

Qy     630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db    2606 ATATCCCTGATCTACGGCTGGCAATTGACACTTTTACTCCTAGCAATTGTTCCCATCATT 2665

Qy     650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db    2666 GCTATAGCAGGAGTGGTTGAAATGAAAATGTTGTCTGGACAAGCGCTGAAAGATAAGAAG 2725

Qy     670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db    2726 GAACTAGAAGGTTCTGGGAAGATCGCTACAGAAGCAATTGAAAACTTTCGCACTGTCGTC 2785

Qy     690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db    2786 TCTTTGACTCGGGAGCAGAAGTTTGAAACTATGTATGCCCAGAGCTTGCAGATACCATAC 2845

Qy     710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db    2846 AGAAATGCTTTGAAGAAAGCGCACGTCTTTGGGATCACTTTCTCCTTCACCCAGGCCATG 2905

Qy     730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db    2906 ATGTATTTCTCCTATGCTGCTTGTTTCCGGTTTGGTGCCTACTTGGTGGCACGAGAACTC 2965

Qy     750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db    2966 ATGACATTTGAAAATGTTCTGTTAGTATTCTCAGCTATTGTCTTTGGTGCCATGGCAGTG 3025

Qy     770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db    3026 GGGCAGGTCAGTTCATTCGCTCCTGACTACGCGAAAGCCAAAGTCTCGGCATCCCACATC 3085

Qy     790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db    3086 ATCAGGATCATTGAGAAAATCCCTGAGATTGACAGCTACAGCACGGAGGGCTTGAAGCCT 3145

Qy     810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
```

```
                 :::      ||||||||::::::|||        |||   |||   ||||||   ||||||
Db      3146 AATATGTTGGAAGGAAATGTGAAATTTAATGGAGTCATGTTCAACTATCCCACCCGACCC 3205

Qy       830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
             ::::::   :::|||:::||||||||||||   :::::::::|||:::|||:::|||
Db      3206 AACATCCCAGTGCTTCAGGGGCTGAGCCTAGAGGTGAAGAAAGGGCAGACGCTGGCCCTC 3265

Qy       850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
             |||||||||||||||||||||||||||||||   ||||||||||||:::|||   ||||||
Db      3266 GTGGGCAGCAGTGGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTTGAGCGCTTCTATGAC 3325

Qy       870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
             |||:::   |||   |||      ||||||   :::   |||:::||||||||||||||||
Db      3326 CCCATGGCCGGAACAGTGTTTCTAGATGGCAAAGAAATAAAGCAACTCAATGTCCAGTGG 3385

Qy       890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
             ||||||:::   :::   ||||||   |||||||||:::||||||:::|||||||||||||
Db      3386 CTCCGCGCCCACCTGGGCATTGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATCGCC 3445

Qy       910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
             ||||||||||||||||||||||||||||||||||||||||   :::||||||   :::|||
Db      3446 GAGAACATTGCCTACGGAGACAACAGCCGTGTCGTGTCTCATGAGGAGATCGTGAAGGCA 3505

Qy       930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
             |||      ||||||||||||   ||||||:::   ||||||||||||||||||||||:::
Db      3506 GCCAAGGAGGCCAACATCCACCAGTTCATCGACTCACTGCCTGAGAAATACAACACCAGA 3565

Qy       950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
             ||||||   ||||||   |||||||||||||||||||||||||||:::||||||||||||
Db      3566 GTGGGAGACAAAGGGACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATCGCCATCGCGCGC 3625

Qy       970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
             ||||||:::::::::|||   |||||||||||||||||||||||||||||||||||||
Db      3626 GCCCTCGTCAGACAGCCTCACATCTTACTTCTGGATGAAGCGACATCAGCTCTGGATACG 3685

Qy       990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
             :::||||||||||||||||||   ||||||||||||||||||   ||||||||||||:::
Db      3686 GAGAGTGAAAAGGTCGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACCTGCATT 3745

Qy      1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
             |||:::   ||||||||||||   |||||||||||||||||||||||||||:::   |||
Db      3746 GTGATCGCGCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATTCAGAAC 3805

Qy      1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
             |||::::::||||||   ||||||||||||:::||||||      :::   |||||||||
Db      3806 GGCCAGGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCCCAGAAAGGCATCTATTTC 3865

Qy      1050 LysLeuValAsnAlaGlnSer 1056
                :::|||:::   |||:::
Db      3866 TCGATGGTCAGTGTGCAGGCT 3886
```

```
RESULT 13
AX105082
LOCUS       AX105082               4279 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 26 from Patent WO0123540.
ACCESSION   AX105082
VERSION     AX105082.1  GI:13921232
KEYWORDS    .
SOURCE      dog.
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 4279)
  AUTHORS   Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and
            Patten,C.J.
  TITLE     P-glycoproteins and uses thereof
  JOURNAL   Patent: WO 0123540-A 26 05-APR-2001;
            GENTEST CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..4279
                     /organism="Canis familiaris"
                     /db_xref="taxon:9615"
     CDS             17..3862
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC37723.1"
                     /db_xref="GI:13921233"
                     /translation="MDPEGGRKGSAEKNFWKMGKKSKKKEKKEKKPTVSTFAMFRYSN
                     WLDRLYMLVGTMAAIIHGAALPLMMLVFGNMTDSFANAGISRNKTFPVIINESITNNT
                     QHFINHLEEMTTYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQILKIRKQFFHAIMRQ
                     EIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSIATFFTGFIVGFTRGWKLTL
                     VILAISPVLGLSAAIWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFGGQKKEL
                     ERYNKNLEEAKGIGIKKAITANISIGAAFLLIYASYALAFWYGTSLVLSSEYTIGQVL
                     TVFFSVLIGAFSIGQASPSIEAFANARGAAYEIFKIIDNKPSIDSYSKSGHKPDNIKG
                     NLEFKNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQRLYDPTD
                     GMVCIDGQDIRTINVRHLREITGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKE
                     ANAYDFIMKLPNKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTE
                     SEAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELMKEKGIY
                     FKLVTMQTRGNEIELENATGESKSESDALEMSPKDSGSSLIKRRSTRRSIHAPQGQDR
                     KLGTKEDLNENVPPVSFWRILKLNSTEWPYFVVGIFCAIINGGLQPAFSIIFSRIIGI
                     FTRDEDPETKRQNSNMFSVLFLVLGIISFITFFLQGFTFGKAGEILTKRLRYMVFRSM
                     LRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAVITQNIANLGTGIIISLIYG
                     WQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRTVVSLTR
                     EQKFEYMYAQSLQVPYRNSLRKAHIFGVSFSITQAMMYFSYAGCFRFGAYLVANEFMN
                     FQDVLLVFSAIVFGAMAVGQVSSFAPDYAKAKVSAAHVIMIIEKSPLIDSYSPHGLKP
                     NTLEGNVTFNEVVFNYPTRPDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERF
                     YDPLAGSVLIDGKEIKHLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSHEE
                     IVQAAKEANIHHFIETLPEKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEA
                     TSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGKVKEHGTHQQL
                     LAQKGIYFSMVSVQAGAKR"
BASE COUNT     1296 a    833 c   1009 g   1141 t
ORIGIN

Alignment Scores:
Pred. No.:              7.17e-195      Length:          4279
Score:                  3115.50        Matches:         602
Percent Similarity:     75.02%         Conservative:    200
Best Local Similarity:  56.31%         Mismatches:      220
Query Match:            58.45%         Indels:          47
DB:                     6              Gaps:            6

US-09-873-409-4 (1-1058) x AX105082 (1-4279)
```

```
Qy         4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
                |||   |||   ||||||||   ||||||:::||||||||||||||||||||||||
Db       734 TCATTTACTGATAAAGAACTCTTGGCCTATGCAAAAGCTGGAGCAGTAGCTGAAGAAGTC 793

Qy        24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
             |||::::::|||||||||||||||||||||            |||:::|||||||||:::
Db       794 TTAGCAGCAATCAGAACTGTGATTGCCTTTGGAGGACAAAAGAAAGAACTTGAA------ 847

Qy        44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63

Db       847 ------------------------------------------------------------ 847

Qy        64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
                               ||||||   :::||||||::::::||||||   |||||||||
Db       848 ---------------AGGTACAACAAAAATTTAGAAGAAGCTAAAGGAATTGGGATAAAG 892

Qy        84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
             :::   |||   :::   :::|||:::||||||   :::      :::      :::|||
Db       893 AAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCATCATAT 952

Qy       104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
                |||||||||||||||||||||||||:::|||:::   |||      |||||||||
Db       953 GCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TATACTATT 1006

Qy       124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
             |||   ||||||   |||||||||||||||:::         ::::::   ||||||   |||
Db      1007 GGACAAGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGACAGGCA 1066

Qy       144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
                |||      |||   ||||||   ||||||||||||||:::   ||||||::::::
Db      1067 TCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1126

Qy       164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
```

```
Db   1127 ATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGATAATATT 1186

Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db   1187 AAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAAGTTAAG 1246

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db   1247 ATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGCAGACAGTGGCGCTGGTTGGGAAC 1306

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db   1307 AGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCCACAGAT 1366

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db   1367 GGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTTCGGGAA 1426

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db   1427 ATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAAAACATT 1486

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db   1487 CGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAAT 1546

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db   1547 GCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAGAGAGGG 1606

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db   1607 GCCCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTTCGCAAC 1666

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db   1667 CCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTG 1726

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db   1727 GTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCTCATCGT 1786

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db   1787 TTGTCTACAGTTCGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATTGTGGAG 1846

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   1847 AAAGGAAATCATGATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTCACAATG 1906

Qy    424 GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 443
Db   1907 CAG---ACAAGAGGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAAAGTGAA 1963

Qy    444 ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--------- 460
Db   1964 AGTGATGCCTTGGAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAAAGAAGA 2020

Qy    461 ------------------------------------AspLysAlaGluGluSerThrGln 468
Db   2021 TCAACTCGCAGGAGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACAAAAGAG 2080

Qy    469 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 488
Db   2081 GACTTGAATGAGAATGTACCTCCAGTTTCCTTCTGGAGGATTCTGAAGCTGAACTCAACT 2140

Qy    489 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 508
Db   2141 GAATGGCCTTATTTTGTGGTTGGTATATTTTGTGCTATTATAAACGGAGGCCTGCAACCA 2200

Qy    509 ValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThr 527
Db   2201 GCATTTTCAATAATATTTTCAAGGATTATAGGGATCTTTACCCGAGATGAGGATCCTGAA 2260

Qy    528 ThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 547
Db   2261 ACAAAACGACAGAATAGTAACATGTTTTCTGTATTGTTTCTAGTCCTTGGAATTATTTCT 2320

Qy    548 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 567
Db   2321 TTTATTACATTTTTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTCACTAAG 2380

Qy    568 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 587
Db   2381 CGGCTTCGATACATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTTGATGAC 2440

Qy    588 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 607
Db   2441 CCTAAAAACACCACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAAGTTAAA 2500

Qy    608 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 627
Db   2501 GGGGCTATAGGTTCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGGACAGGC 2560

Qy    628 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 647
Db   2561 ATTATTATATCCTTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATTGTACCC 2620

Qy    648 ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp 667
Db   2621 ATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGAT 2680

Qy    668 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 687
Db   2681 AAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTCCGAACT 2740

Qy    688 IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 707
Db   2741 GTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTGCAAGTA 2800

Qy    708 GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis 727
Db   2801 CCATACAGAAACTCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATCACCCAG 2860

Qy    728 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla 747
Db   2861 GCAATGATGTATTTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTGGCAAAT 2920

Qy    748 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 767
Db   2921 GAGTTCATGAACTTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGTGCCATG 2980

Qy    768 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 787
Db   2981 GCAGTGGGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCAGCAGCC 3040

Qy    788 HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys 807
Db   3041 CACGTCATCATGATCATTGAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCACGGCCTC 3100

Qy    808 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys 827
Db   3101 AAGCCAAATACGTTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTATCCCACT 3160

Qy    828 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 847
Db   3161 CGACCAGACATCCCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAGACGCTG 3220

Qy    848 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 867
Db   3221 GCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAGCGCTTC 3280

Qy    868 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 887
Db   3281 TATGACCCCTTGGCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTGAATGTC 3340
```

```
Qy     888 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 907
           ||||||||||:::   :::  ||||||  |||||||||:::||||||:::||||||
Db    3341 CAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGACTGCAGC 3400

Qy     908 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 927
           ||||||||||||||||||||||||||||||||||||||||||||||      :::||||||
Db    3401 ATTGCCGAGAACATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAGATTGTG 3460

Qy     928 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn 947
           :::||||||      ||||||||||||  |||||||||  ||||||||||||||||||
Db    3461 CAGGCAGCCAAGGAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAATACAAC 3520

Qy     948 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 967
           |||:::||||||   ||||||  |||||||||||||||||||||||||:::||||||
Db    3521 ACCAGAGTAGGAGACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATTGCCATA 3580

Qy     968 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 987
           ||||||||||||:::::::::|||   |||||||||||||||||||||||||||||||||
Db    3581 GCTCGCGCTCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCTACATCAGCTCTG 3640

Qy     988 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1007
           |||   :::||||||||||||||||||  ||||||||||||||||||   |||||||||
Db    3641 GATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACC 3700

Qy    1008 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1027
           |||:::||||:::   ||||||||||||   |||||||||||||||||||||||||||||
Db    3701 TGCATTGTGATCGCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTGGTGTTT 3760

Qy    1028 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1047
           ||||||||||:::||||||  |||||||||||::::||||||      :::    |||
Db    3761 CAGAATGGCAAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGGCCCAGAAAGGCATC 3820

Qy    1048 TyrPheLysLeuValAsnAlaGlnSer 1056
           ||||||  . :::|||:::   |||:::
Db    3821 TATTTTTCCATGGTCAGTGTCCAGGCT 3847
```

```
RESULT 14
AX105057
LOCUS       AX105057                4279 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 1 from Patent WO0123540.
ACCESSION   AX105057
VERSION     AX105057.1  GI:13921209
KEYWORDS    .
SOURCE      dog.
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 4279)
  AUTHORS   Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and
            Patten,C.J.
  TITLE     P-glycoproteins and uses thereof
  JOURNAL   Patent: WO 0123540-A 1 05-APR-2001;
            GENTEST CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..4279
                     /organism="Canis familiaris"
                     /db_xref="taxon:9615"
     CDS             17..3862
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC37718.1"
                     /db_xref="GI:13921210"
                     /translation="MDPEGGRKGSAEKNFWKMGKKSKKNEKKEKKPTVSTFAMFRYSN
                     WLDRLYMLVGTMAAIIHGAALPLMMLVFGNMTDSFANAGISRNKTFPVIINESITNNT
                     QHFINHLEEEMTTYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQILKIRKQFFHAIMRQ
                     EIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFHSIATFFTGFIVGFTRGWKLTL
                     VILAISPVLGLSAAIWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFGGQKKEL
                     ERYNKNLEEAKGIGIKKAITANISIGAAFLLIYASYALAFWYGTSLVLSSEYSIGQVL
                     TVFFSVLIGAFSIGQASPSIEAFANARGAAYEIFKIIDNKPSIDSYSKSGHKPDNIKG
                     NLEFKNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQRLYDPTD
                     GMVCIDGQDIRTINVRHLREITGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKE
                     ANAYDFIMKLPNKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTE
                     SEAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELMKEKGIY
                     FKLVTMQTRGNEIELENATGESKSESDALEMSPKDSGSSLIKRRSTRRSIHAPQGQDR
                     KLGTKEDLNENVPPVSFWRILKLNSTEWPYFVVGIFCAIINGGLQPAFSIIFSRIIGI
                     FTRDEDPETKRQNSNMFSVLFLVLGIISFITFFLQGFTFGKAGEILTKRLRYMVFRSM
                     LRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAVITQNIANLGTGIIISLIYG
                     WQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRTVVSLTR
                     EQKFEYMYAQSLQVPYRNSLRKAHIFGVSFSITQAMMYFSYAGCFRFGAYLVANEFMN
                     FQDVLLVFSAIVFGAMAVGQVSSFAPDYAKAKVSAAHVIMIIEKSPLIDSYSPHGLKP
                     NTLEGNVTFNEVVFNYPTRPDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERF
                     YDPLAGSVLIDGKEIKHLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSHEE
                     IMQAAKEANIHHFIETLPEKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEA
                     TSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGKVKEHGTHQQL
                     LAQKGIYFSMVSVQAGAKR"
BASE COUNT     1294 a    834 c   1008 g   1143 t
ORIGIN

Alignment Scores:
Pred. No.:               1.13e-194       Length:          4279
Score:                   3112.50         Matches:         601
Percent Similarity:      75.02%          Conservative:    201
Best Local Similarity:   56.22%          Mismatches:      220
Query Match:             58.40%          Indels:          47
DB:                      6               Gaps:            6

US-09-873-409-4 (1-1058) x AX105057 (1-4279)
```

```
Qy       4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
           |||   |||   |||||||||  ||||||:::|||||||||||||||||||||||||||
Db     734 TCATTTACTGATAAAGAACTCTTGGCCTATGCAAAAGCTGGAGCAGTAGCTGAAGAAGTC 793

Qy      24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
           |||::::::|||||||||||||||||||||      |||:::|||||||||:::
Db     794 TTAGCAGCAATCAGAACTGTGATTGCCTTTGGAGGACAAAAGAAAGAACTTGAA------ 847

Qy      44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63

Db     847 ------------------------------------------------------------ 847

Qy      64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
                         ||||||   :::||||||::::::||||||      |||||||||
Db     848 --------------AGGTACAACAAAAATTTAGAAGAAGCTAAAGGAATTGGGATAAAG 892

Qy      84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
           :::   |||   :::   :::|||:::||||||   :::         :::   :::|||
Db     893 AAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCATCATAT 952

Qy     104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
              |||||||||||||||||||||||||:::|||:::   |||      |||:::|||
Db     953 GCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TATTCTATT 1006

Qy     124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
           |||   ||||||   |||||||||||||||:::      ::::::   ||||||   |||
Db    1007 GGACAAGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGACAGGCA 1066

Qy     144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
              |||      |||   ||||||   ||||||||||||||:::   ||||||::::::
Db    1067 TCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1126

Qy     164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
           ||||||   |||||||||||||::::::|||      :::|||:::|||::::::|||
Db    1127 ATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGATAATATT 1186

Qy     184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
           :::|||   :::||||||||||||   |||:::||||||||||||      :::|||
Db    1187 AAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAAGTTAAG 1246

Qy     204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
           ||||||||||||||||||||||||:::::::::||||||:::|||||||||||||||||
Db    1247 ATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTTGGGAAC 1306

Qy     224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
           :::|||   ||||||||||| . ||||||||:::||||||||||||||||||   |||
Db    1307 AGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCCACAGAT 1366
```

```
Qy     244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db    1367 GGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTTCGGGAA 1426

Qy     264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db    1427 ATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAAAACATT 1486

Qy     284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db    1487 CGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAAT 1546

Qy     304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db    1547 GCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAGAGAGGG 1606

Qy     324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db    1607 GCCCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTTCGCAAC 1666

Qy     344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db    1667 CCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTG 1726

Qy     364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db    1727 GTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCTCATCGT 1786

Qy     384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db    1787 TTGTCTACAGTTCGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATTGTGGAG 1846

Qy     404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db    1847 AAAGGAAATCATGATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTCACAATG 1906

Qy     424 GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 443
Db    1907 CAG---ACAAGAGGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAAAGTGAA 1963

Qy     444 ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--------- 460
Db    1964 AGTGATGCCTTGGAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAAAGAAGA 2020

Qy     461 -----------------------------------AspLysAlaGluGluSerThrGln 468
Db    2021 TCAACTCGCAGGAGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACAAAAGAG 2080

Qy     469 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 488
Db    2081 GACTTGAATGAGAATGTACCTCCAGTTTCCTTCTGGAGGATTCTGAAGCTGAACTCAACT 2140

Qy     489 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 508
Db    2141 GAATGGCCTTATTTTGTGGTTGGTATATTTTGTGCTATTATAAACGGAGGCCTGCAACCA 2200

Qy     509 ValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThr 527
Db    2201 GCATTTTCAATAATATTTTCAAGGATTATAGGGATCTTTACCCGAGATGAGGATCCTGAA 2260

Qy     528 ThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 547
Db    2261 ACAAAACGACAGAATAGTAACATGTTTTCTGTATTGTTTCTAGTCCTTGGAATTATTTCT 2320

Qy     548 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 567
Db    2321 TTTATTACATTTTTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTCACTAAG 2380

Qy     568 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 587
Db    2381 CGGCTTCGATACATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTTGATGAC 2440

Qy     588 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 607
Db    2441 CCTAAAAACACCACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAAGTTAAA 2500

Qy     608 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 627
Db    2501 GGGGCTATAGGTTCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGGACAGGC 2560

Qy     628 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 647
Db    2561 ATTATTATATCCTTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATTGTACCC 2620

Qy     648 ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp 667
Db    2621 ATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGAT 2680

Qy     668 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 687
Db    2681 AAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTCCGAACT 2740

Qy     688 IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 707
Db    2741 GTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTGCAAGTA 2800

Qy     708 GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis 727
Db    2801 CCATACAGAAACTCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATCACCCAG 2860

Qy     728 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla 747
Db    2861 GCAATGATGTATTTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTGGCAAAT 2920

Qy     748 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 767
Db    2921 GAGTTCATGAACTTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGTGCCATG 2980

Qy     768 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 787
Db    2981 GCAGTGGGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCAGCAGCC 3040

Qy     788 HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys 807
Db    3041 CACGTCATCATGATCATTGAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCACGGCCTC 3100

Qy     808 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys 827
Db    3101 AAGCCAAATACGTTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTATCCCACT 3160

Qy     828 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 847
Db    3161 CGACCAGACATCCCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAGACGCTG 3220

Qy     848 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 867
Db    3221 GCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAGCGCTTC 3280

Qy     868 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 887
Db    3281 TATGACCCCTTGGCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTGAATGTC 3340

Qy     888 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 907
Db    3341 CAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGACTGCAGC 3400

Qy     908 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 927
Db    3401 ATTGCCGAGAACATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAGATTATG 3460

Qy     928 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn 947
Db    3461 CAGGCAGCCAAGGAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAATACAAC 3520

Qy     948 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 967
```

```
              |||:::||||||  ||||||  ||||||||||||||||||||||||||:::||||||
Db     3521 ACCAGAGTAGGAGACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATTGCCATA 3580

Qy      968 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 987
            ||||||||||||:::::::|||  |||||||||||||||||||||||||||||||||||||
Db     3581 GCTCGCGCTCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCTACATCAGCTCTG 3640

Qy      988 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1007
            |||   :::||||||||||||||||||  ||||||||||||||||    |||||||||
Db     3641 GATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACC 3700

Qy     1008 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1027
            |||:::|||:::   |||||||||||  ||||||||||||||||||||||||
Db     3701 TGCATTGTGATCGCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTGGTGTTT 3760

Qy     1028 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1047
               |||||||||:::||||||  ||||||||||||:::||||||      :::   |||
Db     3761 CAGAATGGCAAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGGCCCAGAAAGGCATC 3820

Qy     1048 TyrPheLysLeuValAsnAlaGlnSer 1056
            ||||||   :::|||:::   |||:::
Db     3821 TATTTTTCCATGGTCAGTGTCCAGGCT 3847

RESULT 15
AX105078
LOCUS       AX105078                4279 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 22 from Patent WO0123540.
ACCESSION   AX105078
VERSION     AX105078.1  GI:13921228
KEYWORDS    .
SOURCE      dog.
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 4279)
  AUTHORS   Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and
            Patten,C.J.
  TITLE     P-glycoproteins and uses thereof
  JOURNAL   Patent: WO 0123540-A 22 05-APR-2001;
            GENTEST CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..4279
                     /organism="Canis familiaris"
                     /db_xref="taxon:9615"
     CDS             17..3862
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC37721.1"
                     /db_xref="GI:13921229"
                     /translation="MDPEGGRKGSAEKNFWKMGKKSKKNEKKEKKPTVSTFAMFRYSN
                     WLDRLYMLVGTMAAIIHGAALPLMMLVFGNMTDSFANAGISRNKTFPVIINESITNNT
                     QHFINHLEEEMTTYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQILKIRKQFFHAIMRQ
                     EIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSIATFFTGFIVGFTRGWKLTL
                     VILAISPVLGLSAAIWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFGGQKKEL
                     ERYNKNLEEAKGIGIKKAITANISIGAAFLLIYASYALAFWYGTSLVLSSEYSIGQVL
                     TVFFSVLIGAFSIGQASPSIEAFANARGAAYEIFKIIDNKPSIDSYSKSGHKPDNIKG
                     NLEFKNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQRLYDPTD
                     GMVCIDGQDIRTINVRHLREITGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKE
                     ANAYDFIMKLPNKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTE
                     SEAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELMKEKGIY
                     FKLVTMQTRGNEIELENATGESKSESDALEMSPKDSGSSLIKRRSTRRSIHAPQGQDR
                     KLGTKEDLNENVPPVSFWRILKLNSTEWPYFVVGIFCAIINGGLQPAFSIIFSRIIGI
                     FTRDEDPETKRQNSNMFSVLFLVLGIISFITFFLQGFTFGKAGEILTKRLRYMVFRSM
                     LRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAVITQNIANLGTGIIISLIYG
                     WQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRTVVSLTR
                     EQKFEYMYAQSLQVPYRNSLRKAHIFGVSFSITQAMMYFSYAGCFRFGAYLVANEFMN
                     FQDVLLVFSAIVFGAMAVGQVSSFAPDYAKAKVSAAHVIMIIEKSPLIDSYSPHGLKP
                     NTLEGNVTFNEVVFNYPTRPDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERF
                     YDPLAGSVLIDGKEIKHLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSHEE
                     IMQAAKEANIHHFIETLPEKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEA
                     TSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGKVKEHGTHQQL
                     LAQKGIYFSMVSVQAGAKR"
BASE COUNT     1295 a    833 c   1008 g   1143 t
ORIGIN

Alignment Scores:
Pred. No.:              1.13e-194      Length:          4279
Score:                  3112.50        Matches:         601
Percent Similarity:     75.02%         Conservative:    201
Best Local Similarity:  56.22%         Mismatches:      220
Query Match:            58.40%         Indels:          47
DB:                     6              Gaps:            6

US-09-873-409-4 (1-1058) x AX105078 (1-4279)

Qy        4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
            |||   |||   ||||||||   ||||||:::||||||||||||||||||||||||
Db      734 TCATTTACTGATAAAGAACTCTTGGCCTATGCAAAAGCTGGAGCAGTAGCTGAAGAAGTC 793

Qy       24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
            |||::::::||||||||||||||||||||||||      |||:::||||||||||:::
Db      794 TTAGCAGCAATCAGAACTGTGATTGCCTTTGGAGGACAAAAGAAAGAACTTGAA------ 847

Qy       44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63

Db      847 ------------------------------------------------------------ 847

Qy       64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
                         ||||||   :::||||||::::::||||||      |||||||||
Db      848 --------------AGGTACAACAAAAATTTAGAAGAAGCTAAAGGAATTGGGATAAAG 892

Qy       84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
            :::   |||   :::   :::|||:::||||||   :::      :::      :::|||
Db      893 AAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCATCATAT 952

Qy      104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
               |||||||||||||||||||||||||||:::|||:::   |||      |||:::|||
Db      953 GCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TATTCTATT 1006

Qy      124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
            |||   ||||||   ||||||||||||||:::      ::::::   ||||||   |||
Db     1007 GGACAAGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGACAGGCA 1066

Qy      144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
               |||      |||   ||||||   |||||||||||||||:::   ||||||::::::
Db     1067 TCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1126

Qy      164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
            ||||||   ||||||||||||||||:::::::|||   :::|||:::||||||:::::::|||
Db     1127 ATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGATAATATT 1186

Qy      184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
            :::|||   :::|||||||||||||||   |||:::||||||||||||      :::|||
Db     1187 AAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAAGTTAAG 1246

Qy      204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
            |||||||||||||||||||||||:::::::::||||||:::|||||||||||||||||
Db     1247 ATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTTGGGAAC 1306

Qy      224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
            :::|||   ||||||||||||   ||||||||:::||||||||||||||||||   |||
Db     1307 AGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCCACAGAT 1366

Qy      244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
            |||   :::   :::|||      ||||||||||   :::|||||||||||   |||:::
Db     1367 GGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTTCGGGAA 1426

Qy      264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
                  ||||||||||||||||||||||||||   ||||||||:::      ||||||
Db     1427 ATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAAAACATT 1486

Qy      284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
            :::||||||||||::::::||||||   :::|||:::|||:::|||   :::|||||||||
Db     1487 CGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAAT 1546

Qy      304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
```

```
Db   1547 GCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAGAGAGGG 1606

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db   1607 GCCCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTTCGCAAC 1666

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db   1667 CCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTG 1726

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db   1727 GTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCTCATCGT 1786

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db   1787 TTGTCTACAGTTCGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATTGTGGAG 1846

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   1847 AAAGGAAATCATGATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTCACAATG 1906

Qy    424 GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 443
Db   1907 CAG---ACAAGAGGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAAAGTGAA 1963

Qy    444 ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--------- 460
Db   1964 AGTGATGCCTTGGAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAAAGAAGA 2020

Qy    461 ------------------------------------AspLysAlaGluGluSerThrGln 468
Db   2021 TCAACTCGCAGGAGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACAAAAGAG 2080

Qy    469 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 488
Db   2081 GACTTGAATGAGAATGTACCTCCAGTTTCCTTCTGGAGGATTCTGAAGCTGAACTCAACT 2140

Qy    489 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 508
Db   2141 GAATGGCCTTATTTTGTGGTTGGTATATTTTGTGCTATTATAAACGGAGGCCTGCAACCA 2200

Qy    509 ValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThr 527
Db   2201 GCATTTTCAATAATATTTTCAAGGATTATAGGGATCTTTACCCGAGATGAGGATCCTGAA 2260

Qy    528 ThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 547
Db   2261 ACAAAACGACAGAATAGTAACATGTTTTCTGTATTGTTTCTAGTCCTTGGAATTATTTCT 2320

Qy    548 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 567
Db   2321 TTTATTACATTTTTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTCACTAAG 2380

Qy    568 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 587
Db   2381 CGGCTTCGATACATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTTGATGAC 2440

Qy    588 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 607
Db   2441 CCTAAAAACACCACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAAGTTAAA 2500

Qy    608 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 627
Db   2501 GGGGCTATAGGTTCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGGACAGGC 2560

Qy    628 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 647
Db   2561 ATTATTATATCCTTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATTGTACCC 2620

Qy    648 ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp 667
Db   2621 ATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGAT 2680

Qy    668 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 687
Db   2681 AAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTCCGAACT 2740

Qy    688 IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 707
Db   2741 GTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTGCAAGTA 2800

Qy    708 GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis 727
Db   2801 CCATACAGAAACTCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATCACCCAG 2860

Qy    728 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla 747
Db   2861 GCAATGATGTATTTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTGGCAAAT 2920

Qy    748 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 767
Db   2921 GAGTTCATGAACTTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGTGCCATG 2980

Qy    768 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 787
Db   2981 GCAGTGGGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCAGCAGCC 3040

Qy    788 HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys 807
Db   3041 CACGTCATCATGATCATTGAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCACGGCCTC 3100

Qy    808 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys 827
Db   3101 AAGCCAAATACGTTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTATCCCACT 3160

Qy    828 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 847
Db   3161 CGACCAGACATCCCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAGACGCTG 3220

Qy    848 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 867
Db   3221 GCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAGCGCTTC 3280

Qy    868 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 887
Db   3281 TATGACCCCTTGGCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTGAATGTC 3340

Qy    888 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 907
Db   3341 CAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGACTGCAGC 3400

Qy    908 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 927
Db   3401 ATTGCCGAGAACATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAGATTATG 3460

Qy    928 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn 947
Db   3461 CAGGCAGCCAAGGAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAATACAAC 3520

Qy    948 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 967
Db   3521 ACCAGAGTAGGAGACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATTGCCATA 3580

Qy    968 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 987
Db   3581 GCTCGCGCTCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCTACATCAGCTCTG 3640

Qy    988 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1007
Db   3641 GATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACC 3700

Qy   1008 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1027
Db   3701 TGCATTGTGATCGCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTGGTGTTT 3760
```

```
Qy    1028 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1047
              ||||||||:::||||||   |||||||||||:::||||||       :::    |||
Db    3761 CAGAATGGCAAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGGCCCAGAAAGGCATC 3820

Qy    1048 TyrPheLysLeuValAsnAlaGlnSer 1056
           ||||||   :::|||:::   |||:::
Db    3821 TATTTTTCCATGGTCAGTGTCCAGGCT 3847

Search completed: March 31, 2003, 02:13:48
Job time : 14525 secs
```

GenCore version 5.1.4_p5_4578


```
OM protein - nucleic search, using frame_plus_p2n model

Run on:          March 30, 2003, 01:20:01 ; Search time 1034.02 Seconds
                                            (without alignments)
                                            2304.231 Million cell updates/sec

Title:           US-09-873-409-4
Perfect score:   5330
Sequence:        1 MVISLTSKELSAYSKAGAVA..........QELLRNRDIYFKLVNAQSVQ 1058

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters:      4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115416_19216/app_query.fasta_1.7544
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_4247_@runat_27032003_115416_19216 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       N_Geneseq_101002:*
                 1:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
                 2:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
                 3:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
                 4:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
                 5:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
                 6:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
                 7:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
                 8:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
                 9:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
                 10:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
                 11:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
                 12:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
                 13:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
                 14:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
                 15:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
                 16:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
                 17:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
                 18:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
                 19:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
                 20:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
                 21:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
                 22:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
                 23:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
                 24:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	4985.5	93.5	3699	24	ABK83223	Human transporter
2	3115.5	58.5	4279	22	AAD03506	Dog P-glycoprotein
3	3112.5	58.4	4279	22	AAD03488	Dog P-glycoprotein
4	3112.5	58.4	4279	22	AAD03504	Dog P-glycoprotein
5	3112.5	58.4	4279	22	AAD03505	Dog P-glycoprotein
6	3111.5	58.4	4317	22	AAD03489	Dog P-glycoprotein
7	3106.5	58.3	4369	21	AAZ52047	Rat multidrug resi
8	3106.5	58.3	4425	21	AAZ52048	Rat multidrug resi
9	3100.5	58.2	3860	21	AAZ49332	Human wild-type mu
10	3100.5	58.2	3860	21	AAZ49333	Human G185V mutant
11	3100.5	58.2	3860	24	ABA94365	Human BCRP DNA rel
12	3100.5	58.2	3860	24	ABA94366	Human BCRP DNA rel
13	3100.5	58.2	4186	22	AAF86127	Cynomologous monke
14	3100.5	58.2	4195	22	AAF86128	Cynomologous monke
15	3100.5	58.2	4264	19	AAV66533	Mutated human P-gl
16	3100.5	58.2	4264	19	AAV66534	Mutated human P-gl
17	3100.5	58.2	4349	22	AAH57442	Human intestine ce
18	3100.5	58.2	4646	21	AAZ94738	Human ATP binding
19	3100.5	58.2	4646	24	AAD38994	Human mdr1 gene.
20	3100.5	58.2	4646	24	ABL68592	Kidney cancer rela
21	3100.5	58.2	4646	24	ABL68880	Kidney cancer rela
22	3100.5	58.2	4669	8	AAN70752	Sequence of human
23	3100.5	58.2	6505	17	AAT13394	Hybrid vector pSF-
24	3100.5	58.2	8630	21	AAZ24041	Retroviral M4 mdr-
25	3100.5	58.2	8630	21	AAZ24042	Retroviral vector
26	3097.5	58.1	4669	14	AAQ52726	Sequence of human
27	3090.5	58.0	4788	21	AAZ49335	Murine multidrug r
28	3090.5	58.0	4788	24	ABA94368	Mouse BCRP DNA rel
29	3089.5	58.0	4646	15	AAQ72872	Human multidrug re
30	3086	57.9	4233	21	AAZ90198	Rat mdr1b2 (multis
31	3086	57.9	4233	22	AAF27498	Rat mdr1b2 multidr
32	3085.5	57.9	4378	11	AAQ04522	Multidrug Resistan
33	3085	57.9	3924	21	AAZ94742	Human ATP binding
34	3085	57.9	3924	21	AAZ88974	Human MDR-3 DNA.
35	3085	57.9	3924	24	ABN95801	Gene #2299 used to
36	3085	57.9	4189	21	AAZ49334	Murine multidrug r
37	3085	57.9	4189	24	ABA94367	Mouse BCRP DNA rel
38	3085	57.9	4313	14	AAQ38950	Mouse multidrug re
39	3084	57.9	3840	24	ABL91687	Human polynucleoti
40	3084	57.9	3988	21	AAZ88973	Human MDR-1 DNA.
41	3073.5	57.7	4669	19	AAV32645	Human P glycoprote
42	3073.5	57.7	4669	24	ABK52041	cDNA encoding huma
43	3044.5	57.1	4254	24	ABK63517	Rat sequence diffe
44	3043	57.1	3912	24	ABK63653	Rat sequence diffe
45	2515.5	47.2	4776	21	AAZ94744	Human ATP binding

ALIGNMENTS

```
RESULT 1
ABK83223
ID   ABK83223 standard; cDNA; 3699 BP.
XX
AC   ABK83223;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human transporter and ion channel, TRICH14, Incyte ID 7472030CB1, cDNA.
XX
KW   Human; ss; gene; transporter and ion channel; TRICH; transport disorder;
KW   neurological disorder; muscle disorder; immunological disorder; cancer;
KW   scleroderma; systemic lupus erythematosus; allergy; leukaemia;
KW   cell proliferative disorder; cervical cancer; breast cancer;
KW   neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;
KW   myotonic dystrophy; catatonia; endocrine disorder; diabetes;
KW   Grave's disease; gastrointestinal disorder; Crohn's disease;
KW   renal disorder; Good pasture's syndrome; viral infection; cirrhosis;
```

```
KW   bacterial infection; fungal infection; parasitic infection;
KW   protozoal infection; helminthic infection; cardiovascular disorder;
KW   atherosclerosis; hepatic disease.
XX
OS   Homo sapiens.
XX
PN   WO200240541-A2.
XX
PD   23-MAY-2002.
XX
PF   25-OCT-2001; 2001WO-US46055.
XX
PR   27-OCT-2000; 2000US-243989P.
PR   03-NOV-2000; 2000US-245904P.
PR   09-NOV-2000; 2000US-247673P.
PR   17-NOV-2000; 2000US-249661P.
PR   20-NOV-2000; 2000US-252232P.
PR   01-DEC-2000; 2000US-250790P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Tang YT,  Yue H,  Nguyen DB,  Hafalia AJA,  Elliott VS,  Lu Y;
PI   Walia NK,  Yao MG,  Baughn MR,  Gandhi AR,  Ding L,  Sanjanwala M;
PI   Ramkumar J,  Arvizu C,  Gietzen KJ,  Lal PG,  Azimzai Y,  Khan FA;
PI   Thangavelu K,  Thornton M,  Lu DAM,  Tribouley CM,  Warren BA;
PI   Ison CH,  Das D,  Raumann BE,  Policky JL,  Kearney L;
XX
DR   WPI; 2002-463570/49.
DR   P-PSDB; ABG61544.
XX
PT   New transporters and ion channels (TRICH) polypeptides, useful for
PT   diagnosing, preventing, and treating disorders associated with an
PT   abnormal expression or activity of TRICH, e.g. immunological, muscular
PT   or renal disorders  -
XX
PS   Claim 5; Page 172; 178pp; English.
XX
CC   The invention relates to human transporters and ion channels (TRICH)
CC   polypeptides, a naturally occurring amino acid sequence 90 % identical to
CC   TRICH, a biologically active fragment of TRICH or an immunogenic fragment
CC   of TRICH. Also included are an isolated polynucleotide encoding TRICH,
CC   a recombinant polynucleotide comprising a promoter sequence operably
CC   linked to the TRICH polynucleotide, a cell transformed with the
CC   recombinant polynucleotide, a transgenic organism comprising the
CC   recombinant polynucleotide, an isolated antibody that binds specifically
CC   to TRICH, and screening for compounds which bind to TRICH, modulate
CC   TRICH, modulate TRICH expression or are ant/agonists of TRICH.
CC   The polypeptides are useful for diagnosing, treating, and
CC   preventing transport, neurological, muscle, immunological disorders
CC   (e.g. scleroderma, systemic lupus erythematosus, allergies), cell
CC   proliferative disorders such as cancers (e.g. leukaemia, cervical or
CC   breast cancers), neurodegenerative disorders (e.g. Parkinson's disease,
CC   Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy,
CC   catatonia), endocrine disorders (e.g. diabetes, Grave's disease),
CC   gastrointestinal disorders (e.g. Crohn's disease), renal disorders
CC   (e.g. Good pasture's syndrome), viral, bacterial, fungal, parasitic,
CC   protozoal and helminthic infections, cardiovascular disorders (e.g.
CC   atherosclerosis), or hepatic diseases (e.g. cirrhosis) and many
CC   other diseases and disorders detailed in the specification. They can also
CC   be used in assessing the effects of exogenous compounds on the
CC   expression of nucleic acid and amino acid sequences of transporters and
CC   ion channels. TRICH or its fragments may also be used in screening for
CC   compounds that specifically bind to and modulate the activity of TRICH.
CC   The polynucleotides can be used to create knock-in humanised animals or
CC   transgenic animals to model human disease. The present sequence
CC   encodes a TRICH protein.
XX
SQ   Sequence 3699 BP; 1116 A; 707 C; 860 G; 1016 T; 0 other;

Alignment Scores:
Pred. No.:                0            Length:        3699
Score:                    4985.50      Matches:       1002
Percent Similarity:       94.80%       Conservative:  1

Best Local Similarity:   94.71%        Mismatches:    2
Query Match:             93.54%        Indels:        53
DB:                      24            Gaps:          2

US-09-873-409-4 (1-1058) x ABK83223 (1-3699)

Qy       1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     682 ATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCA 741

Qy      21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     742 GAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAGGAGAAAGAACTT 801

Qy      41 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60
           |||
Db     802 CAA--------------------------------------------------------- 804

Qy      61 LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80
                                  |||||||||||||||||||||||||||||||||||||
Db     805 -----------------------AGGTATACACAGAATCTCAAAGATGCAAAGGATTTT 840

Qy      81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 GGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAAT 900

Qy     101 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGA 960

Qy     121 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 TATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGTAGTTATTGCATT 1020

Qy     141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATT 1080

Qy     161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180
           ||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db    1081 TTCCAGGTTATTGATAAGAAACCCAGTATAGGTAACTTTTCCACAGCTGGATATAAACCT 1140

Qy     181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 GAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGGCCA 1200

Qy     201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
           ||||||||||||||||||||||||||||||   ||||||||||||||||||||||||||
Db    1201 TCTATCAAGATTCTGAAAGGTCTGAATCTCGGAATTAAGTCTGGAGAGACAGTCGCCTTG 1260

Qy     221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 1320

Qy     241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 CCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCAT 1380

Qy     261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 TATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGT 1440

Qy     281 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 AACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGG 1500

Qy     301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 GAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGG 1560

Qy     321 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 GAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTA 1620
```

```
Qy    341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGC 1680

Qy    361 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 AAGTCAGCTGTTCAAGCTGCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTA 1740

Qy    381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GCACACCGACTTTCTACTATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATG 1800

Qy    401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 CTGGCGGAGAAAGGAGCGCATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTT 1860

Qy    421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACT 1920

Qy    441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 GAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATT 1980

Qy    461 AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 GACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTA 2040

Qy    481 LysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 AAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCT 2100

Qy    501 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 GTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATG 2160

Qy    521 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 TTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTC 2220

Qy    541 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 560
          |||||||||||||||||||||||||||||||||||||||
Db   2221 GTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATG--------------------- 2259

Qy    561 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 580
                                                                   |||
Db   2260 ---------------------------------------------------------CAG 2262

Qy    581 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2263 GATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCC 2322

Qy    601 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2323 ATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAAT 2382

Qy    621 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2383 GCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTC 2442

Qy    641 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2443 CTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATG 2502

Qy    661 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2503 ACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAA 2562

Qy    681 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2563 GCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATG 2622

Qy    701 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2623 TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGA 2682

Qy    721 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2683 AGCTGTTATGCATTCAGCCATGCCTTTATATATTTGCCTATGCGGCAGGGTTTCGATTT 2742

Qy    741 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2743 GGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACT 2802

Qy    761 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 780
          ||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db   2803 GCAATTGCATATGGAGCTATGGCCATCGGAGAAACGCTCGTTTTGGCTCCTGAATATTCC 2862

Qy    781 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2863 AAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGAC 2922

Qy    801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2923 AGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAA 2982

Qy    821 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2983 GTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGT 3042

Qy    841 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3043 ATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCT 3102

Qy    861 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3103 GTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTG 3162

Qy    881 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3163 GATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAG 3222

Qy    901 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3223 CCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTG 3282

Qy    921 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3283 GTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAA 3342

Qy    941 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3343 GGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGC 3402

Qy    961 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3403 CAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTG 3462

Qy    981 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3463 GATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGAT 3522

Qy   1001 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3523 AAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAC 3582

Qy   1021 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3583 GCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAG 3642

Qy   1041 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1058
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3643 CTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3696
```

RESULT 2

```
AAD03506
ID   AAD03506 standard; cDNA; 4279 BP.
XX
AC   AAD03506;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) allelic variant (Genotype D) cDNA.
XX
KW   Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KW   MDR1; drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag=  a
FT                   /product= "Dog P-glycoprotein (PGP) allelic variant
FT                   (Genotype D) protein"
FT   allele          replace (91, T)
FT                   /*tag=  b
FT   allele          replace (607, C)
FT                   /*tag=  c
FT   allele          replace (1001, T)
FT                   /*tag=  c
FT   allele          replace (3458, A)
FT                   /*tag=  c
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00310.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 9; Page 102-107; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) allelic variant
CC   (Genotype D) cDNA. The PGP enzyme functions as an efflux pump
CC   exporting small molecules across the cell membrane. This enzyme
CC   is a member of the ABC transporter family.
XX
SQ   Sequence 4279 BP; 1296 A; 833 C; 1009 G; 1141 T; 0 other;

Alignment Scores:
Pred. No.:              8.38e-284      Length:         4279
Score:                  3115.50        Matches:        602
Percent Similarity:     75.02%         Conservative:   200
```

```
Best Local Similarity:  56.31%      Mismatches:    220
Query Match:            58.45%      Indels:        47
DB:                     22          Gaps:          6

US-09-873-409-4 (1-1058) x AAD03506 (1-4279)

Qy      4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
Db    734 TCATTTACTGATAAAGAACTCTTGGCCTATGCAAAAGCTGGAGCAGTAGCTGAAGAAGTC 793

Qy     24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db    794 TTAGCAGCAATCAGAACTGTGATTGCCTTTGGAGGACAAAAGAAAGAACTTGAA------ 847

Qy     44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db    847 ------------------------------------------------------------ 847

Qy     64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db    848 ---------------AGGTACAACAAAAATTTAGAAGAAGCTAAAGGAATTGGGATAAAG 892

Qy     84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db    893 AAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCATCATAT 952

Qy    104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db    953 GCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TATACTATT 1006

Qy    124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db   1007 GGACAAGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGACAGGCA 1066

Qy    144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db   1067 TCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1126

Qy    164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db   1127 ATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGATAATATT 1186

Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db   1187 AAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAAGTTAAG 1246

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db   1247 ATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTTGGGAAC 1306

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db   1307 AGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCCACAGAT 1366

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db   1367 GGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTTCGGGAA 1426

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db   1427 ATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAAAACATT 1486

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db   1487 CGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAAT 1546

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db   1547 GCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAGAGAGGG 1606

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db   1607 GCCCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTTCGCAAC 1666
```

```
Qy     344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db    1667 CCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTG 1726

Qy     364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db    1727 GTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCTCATCGT 1786

Qy     384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db    1787 TTGTCTACAGTTCGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATTGTGGAG 1846

Qy     404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db    1847 AAAGGAAATCATGATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTCACAATG 1906

Qy     424 GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 443
Db    1907 CAG---ACAAGAGGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAAAGTGAA 1963

Qy     444 ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--------- 460
Db    1964 AGTGATGCCTTGGAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAAAGAAGA 2020

Qy     461 -----------------------------------AspLysAlaGluGluSerThrGln 468
Db    2021 TCAACTCGCAGGAGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACAAAAGAG 2080

Qy     469 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 488
Db    2081 GACTTGAATGAGAATGTACCTCCAGTTTCCTTCTGGAGGATTCTGAAGCTGAACTCAACT 2140

Qy     489 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 508
Db    2141 GAATGGCCTTATTTTGTGGTTGGTATATTTTGTGCTATTATAAACGGAGGCCTGCAACCA 2200

Qy     509 ValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThr 527
Db    2201 GCATTTTCAATAATATTTTCAAGGATTATAGGGATCTTTACCCGAGATGAGGATCCTGAA 2260

Qy     528 ThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 547
Db    2261 ACAAAACGACAGAATAGTAACATGTTTTCTGTATTGTTTCTAGTCCTTGGAATTATTTCT 2320

Qy     548 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 567
Db    2321 TTTATTACATTTTTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTCACTAAG 2380

Qy     568 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 587
Db    2381 CGGCTTCGATACATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTTGATGAC 2440

Qy     588 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 607
Db    2441 CCTAAAAACACCACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAAGTTAAA 2500

Qy     608 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 627
Db    2501 GGGGCTATAGGTTCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGGACAGGC 2560

Qy     628 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 647
Db    2561 ATTATTATATCCTTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATTGTACCC 2620

Qy     648 ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp 667
Db    2621 ATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGAT 2680

Qy     668 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 687
Db    2681 AAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTCCGAACT 2740

Qy     688 IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 707
Db    2741 GTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTGCAAGTA 2800

Qy     708 GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis 727
Db    2801 CCATACAGAAACTCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATCACCCAG 2860

Qy     728 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla 747
Db    2861 GCAATGATGTATTTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTGGCAAAT 2920

Qy     748 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 767
Db    2921 GAGTTCATGAACTTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGTGCCATG 2980

Qy     768 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 787
Db    2981 GCAGTGGGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCAGCAGCC 3040

Qy     788 HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys 807
Db    3041 CACGTCATCATGATCATTGAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCACGGCCTC 3100

Qy     808 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys 827
Db    3101 AAGCCAAATACGTTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTATCCCACT 3160

Qy     828 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 847
Db    3161 CGACCAGACATCCCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAGACGCTG 3220

Qy     848 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 867
Db    3221 GCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAGCGCTTC 3280

Qy     868 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 887
Db    3281 TATGACCCCTTGGCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTGAATGTC 3340

Qy     888 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 907
Db    3341 CAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGACTGCAGC 3400

Qy     908 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 927
Db    3401 ATTGCCGAGAACATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAGATTGTG 3460

Qy     928 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn 947
Db    3461 CAGGCAGCCAAGGAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAATACAAC 3520

Qy     948 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 967
Db    3521 ACCAGAGTAGGAGACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATTGCCATA 3580

Qy     968 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 987
Db    3581 GCTCGCGCTCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCTACATCAGCTCTG 3640

Qy     988 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1007
Db    3641 GATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACC 3700

Qy    1008 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1027
Db    3701 TGCATTGTGATCGCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTGGTGTTT 3760

Qy    1028 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1047
Db    3761 CAGAATGGCAAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGGCCCAGAAAGGCATC 3820

Qy    1048 TyrPheLysLeuValAsnAlaGlnSer 1056
```

```
            |||||    :::|||:::    |||:::
Db    3821 TATTTTTCCATGGTCAGTGTCCAGGCT 3847

RESULT 3
AAD03488
ID   AAD03488 standard; cDNA; 4279 BP.
XX
AC   AAD03488;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) cDNA #1.
XX
KW   Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag=  a
FT                   /product= "Dog P-glycoprotein (PGP) #1"
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;    99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00303.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 3; Page 58-63; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) cDNA. This
CC   sequence is also referred as Genotype C cDNA. The
CC   PGP enzyme functions as an efflux pump exporting small molecules
CC   across the cell membrane. This enzyme is a member of the ABC
CC   transporter family.
XX
SQ   Sequence 4279 BP; 1294 A; 834 C; 1008 G; 1143 T; 0 other;

Alignment Scores:
Pred. No.:              1.61e-283     Length:        4279
Score:                  3112.50       Matches:       601
Percent Similarity:     75.02%        Conservative:  201
Best Local Similarity:  56.22%        Mismatches:    220
Query Match:            58.40%        Indels:        47
DB:                     22            Gaps:          6
```

```
US-09-873-409-4 (1-1058) x AAD03488 (1-4279)

Qy      4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
Db    734 TCATTTACTGATAAAGAACTCTTGGCCTATGCAAAAGCTGGAGCAGTAGCTGAAGAAGTC 793

Qy     24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db    794 TTAGCAGCAATCAGAACTGTGATTGCCTTTGGAGGACAAAAGAAAGAACTTGAA------ 847

Qy     44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db    847 ------------------------------------------------------------ 847

Qy     64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db    848 --------------AGGTACAACAAAAATTTAGAAGAAGCTAAAGGAATTGGGATAAAG 892

Qy     84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db    893 AAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCATCATAT 952

Qy    104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db    953 GCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TATTCTATT 1006

Qy    124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db   1007 GGACAAGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTAGTATTGGACAGGCA 1066

Qy    144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db   1067 TCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1126

Qy    164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db   1127 ATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGATAATATT 1186

Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db   1187 AAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAAGTTAAG 1246

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db   1247 ATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTTGGGAAC 1306

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db   1307 AGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCCACAGAT 1366

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db   1367 GGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTTCGGGAA 1426

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db   1427 ATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAAAACATT 1486

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db   1487 CGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAAT 1546

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db   1547 GCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAGAGAGGG 1606

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db   1607 GCCCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTTCGCAAC 1666

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db   1667 CCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTG 1726
```

```
Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db   1727 GTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCTCATCGT 1786

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db   1787 TTGTCTACAGTTCGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATTGTGGAG 1846

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   1847 AAAGGAAATCATGATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTCACAATG 1906

Qy    424 GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 443
Db   1907 CAG---ACAAGAGGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAAAGTGAA 1963

Qy    444 ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--------- 460
Db   1964 AGTGATGCCTTGGAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAAAGAAGA 2020

Qy    461 ------------------------------------AspLysAlaGluGluSerThrGln 468
Db   2021 TCAACTCGCAGGAGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACAAAAGAG 2080

Qy    469 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 488
Db   2081 GACTTGAATGAGAATGTACCTCCAGTTTCCTTCTGGAGGATTCTGAAGCTGAACTCAACT 2140

Qy    489 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 508
Db   2141 GAATGGCCTTATTTTGTGGTTGGTATATTTTGTGCTATTATAAACGGAGGCCTGCAACCA 2200

Qy    509 ValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThr 527
Db   2201 GCATTTTCAATAATATTTTCAAGGATTATAGGGATCTTTACCCGAGATGAGGATCCTGAA 2260

Qy    528 ThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 547
Db   2261 ACAAAACGACAGAATAGTAACATGTTTTCTGTATTGTTTCTAGTCCTTGGAATTATTTCT 2320

Qy    548 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 567
Db   2321 TTTATTACATTTTTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTCACTAAG 2380

Qy    568 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 587
Db   2381 CGGCTTCGATACATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTTGATGAC 2440

Qy    588 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 607
Db   2441 CCTAAAAACACCACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAAGTTAAA 2500

Qy    608 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 627
Db   2501 GGGGCTATAGGTTCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGGACAGGC 2560

Qy    628 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 647
Db   2561 ATTATTATATCCTTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATTGTACCC 2620

Qy    648 ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp 667
Db   2621 ATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGAT 2680

Qy    668 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 687
Db   2681 AAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTCCGAACT 2740

Qy    688 IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 707
Db   2741 GTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTGCAAGTA 2800

Qy    708 GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis 727
Db   2801 CCATACAGAAACTCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATCACCCAG 2860

Qy    728 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla 747
Db   2861 GCAATGATGTATTTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTGGCAAAT 2920

Qy    748 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 767
Db   2921 GAGTTCATGAACTTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGTGCCATG 2980

Qy    768 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 787
Db   2981 GCAGTGGGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCAGCAGCC 3040

Qy    788 HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys 807
Db   3041 CACGTCATCATGATCATTGAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCACGGCCTC 3100

Qy    808 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys 827
Db   3101 AAGCCAAATACGTTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTATCCCACT 3160

Qy    828 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 847
Db   3161 CGACCAGACATCCCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAGACGCTG 3220

Qy    848 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 867
Db   3221 GCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAGCGCTTC 3280

Qy    868 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 887
Db   3281 TATGACCCCTTGGCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTGAATGTC 3340

Qy    888 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 907
Db   3341 CAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGACTGCAGC 3400

Qy    908 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 927
Db   3401 ATTGCCGAGAACATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAGATTATG 3460

Qy    928 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn 947
Db   3461 CAGGCAGCCAAGGAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAATACAAC 3520

Qy    948 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 967
Db   3521 ACCAGAGTAGGAGACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATTGCCATA 3580

Qy    968 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 987
Db   3581 GCTCGCGCTCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCTACATCAGCTCTG 3640

Qy    988 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1007
Db   3641 GATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACC 3700

Qy   1008 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1027
Db   3701 TGCATTGTGATCGCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTGGTGTTT 3760

Qy   1028 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1047
Db   3761 CAGAATGGCAAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGGCCCAGAAAGGCATC 3820

Qy   1048 TyrPheLysLeuValAsnAlaGlnSer 1056
Db   3821 TATTTTTCCATGGTCAGTGTCCAGGCT 3847
```

RESULT 4

```
AAD03504
ID   AAD03504 standard; cDNA; 4279 BP.
XX
AC   AAD03504;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) allelic variant (Genotype A) cDNA.
XX
KW   Dog; P-glycoprotein allelic variant; multidrug transporter; MDR1;
KW   drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag=  a
FT                   /product= "Dog P-glycoprotein (PGP) allelic variant
FT                   (Genotype A) protein"
FT   allele          replace (607, C)
FT                   /*tag=  b
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00308.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 9; Page 85-90; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) allelic variant
CC   (Genotype A) cDNA. The PGP enzyme functions as an efflux pump
CC   exporting small molecules across the cell membrane. This enzyme
CC   is a member of the ABC transporter family.
XX
SQ   Sequence 4279 BP; 1295 A; 833 C; 1008 G; 1143 T; 0 other;

Alignment Scores:
Pred. No.:              1.61e-283      Length:         4279
Score:                  3112.50        Matches:        601
Percent Similarity:     75.02%         Conservative:   201
Best Local Similarity:  56.22%         Mismatches:     220
Query Match:            58.40%         Indels:         47
DB:                     22             Gaps:           6

US-09-873-409-4 (1-1058) x AAD03504 (1-4279)

Qy      4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
Db    734 TCATTTACTGATAAAGAACTCTTGGCCTATGCAAAAGCTGGAGCAGTAGCTGAAGAAGTC 793

Qy     24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db    794 TTAGCAGCAATCAGAACTGTGATTGCCTTTGGAGGACAAAAGAAAGAACTTGAA------ 847

Qy     44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db    847 ------------------------------------------------------------ 847

Qy     64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db    848 ---------------AGGTACAACAAAAATTTAGAAGAAGCTAAAGGAATTGGGATAAAG 892

Qy     84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db    893 AAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCATCATAT 952

Qy    104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db    953 GCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TATTCTATT 1006

Qy    124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db   1007 GGACAAGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGACAGGCA 1066

Qy    144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db   1067 TCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1126

Qy    164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db   1127 ATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGATAATATT 1186

Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db   1187 AAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAAGTTAAG 1246

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db   1247 ATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTTGGGAAC 1306

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db   1307 AGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCCACAGAT 1366

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db   1367 GGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTTCGGGAA 1426

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db   1427 ATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAAAACATT 1486

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db   1487 CGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAAT 1546

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db   1547 GCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAGAGAGGG 1606

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db   1607 GCCCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTTCGCAAC 1666

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db   1667 CCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTG 1726

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
```

```
Db   1727 GTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCTCATCGT 1786

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db   1787 TTGTCTACAGTTCGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATTGTGGAG 1846

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   1847 AAAGGAAATCATGATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTCACAATG 1906

Qy    424 GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 443
Db   1907 CAG---ACAAGAGGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAAAGTGAA 1963

Qy    444 ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--------- 460
Db   1964 AGTGATGCCTTGGAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAAAGAAGA 2020

Qy    461 -----------------------------------AspLysAlaGluGluSerThrGln 468
Db   2021 TCAACTCGCAGGAGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACAAAAGAG 2080

Qy    469 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 488
Db   2081 GACTTGAATGAGAATGTACCTCCAGTTTCCTTCTGGAGGATTCTGAAGCTGAACTCAACT 2140

Qy    489 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 508
Db   2141 GAATGGCCTTATTTTGTGGTTGGTATATTTGTGCTATTATAAACGGAGGCCTGCAACCA 2200

Qy    509 ValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThr 527
Db   2201 GCATTTTCAATAATATTTTCAAGGATTATAGGGATCTTTACCCGAGATGAGGATCCTGAA 2260

Qy    528 ThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 547
Db   2261 ACAAAACGACAGAATAGTAACATGTTTTCTGTATTGTTTCTAGTCCTTGGAATTATTTCT 2320

Qy    548 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 567
Db   2321 TTTATTACATTTTTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTCACTAAG 2380

Qy    568 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 587
Db   2381 CGGCTTCGATACATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTTGATGAC 2440

Qy    588 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 607
Db   2441 CCTAAAAACACCACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAAGTTAAA 2500

Qy    608 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 627
Db   2501 GGGGCTATAGGTTCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGGACAGGC 2560

Qy    628 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 647
Db   2561 ATTATTATATCCTTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATTGTACCC 2620

Qy    648 ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp 667
Db   2621 ATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGAT 2680

Qy    668 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 687
Db   2681 AAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTCCGAACT 2740

Qy    688 IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 707
Db   2741 GTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTGCAAGTA 2800

Qy    708 GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis 727
Db   2801 CCATACAGAAACTCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATCACCCAG 2860

Qy    728 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla 747
Db   2861 GCAATGATGTATTTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTGGCAAAT 2920

Qy    748 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 767
Db   2921 GAGTTCATGAACTTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGTGCCATG 2980

Qy    768 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 787
Db   2981 GCAGTGGGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCAGCAGCC 3040

Qy    788 HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys 807
Db   3041 CACGTCATCATGATCATTGAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCACGGCCTC 3100

Qy    808 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys 827
Db   3101 AAGCCAAATACGTTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTATCCCACT 3160

Qy    828 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 847
Db   3161 CGACCAGACATCCCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAGACGCTG 3220

Qy    848 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 867
Db   3221 GCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAGCGCTTC 3280

Qy    868 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 887
Db   3281 TATGACCCCTTGGCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTGAATGTC 3340

Qy    888 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 907
Db   3341 CAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGACTGCAGC 3400

Qy    908 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 927
Db   3401 ATTGCCGAGAACATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAGATTATG 3460

Qy    928 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn 947
Db   3461 CAGGCAGCCAAGGAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAATACAAC 3520

Qy    948 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 967
Db   3521 ACCAGAGTAGGAGACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATTGCCATA 3580

Qy    968 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 987
Db   3581 GCTCGCGCTCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCTACATCAGCTCTG 3640

Qy    988 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1007
Db   3641 GATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACC 3700

Qy   1008 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1027
Db   3701 TGCATTGTGATCGCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTGGTGTTT 3760

Qy   1028 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1047
Db   3761 CAGAATGGCAAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGGCCCAGAAAGGCATC 3820

Qy   1048 TyrPheLysLeuValAsnAlaGlnSer 1056
Db   3821 TATTTTTCCATGGTCAGTGTCCAGGCT 3847
```

RESULT 5
AAD03505
ID AAD03505 standard; cDNA; 4279 BP.

```
XX
AC   AAD03505;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.
XX
KW   Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KW   MDR1; drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag=  a
FT                   /product= "Dog P-glycoprotein (PGP) allelic variant
FT                   (Genotype B) protein"
FT   allele          replace (91, T)
FT                   /*tag=  b
FT   allele          replace (607, C)
FT                   /*tag=  c
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00309.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 9; Page 93-99; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) allelic variant
CC   (Genotype B) cDNA. The PGP enzyme functions as an efflux pump
CC   exporting small molecules across the cell membrane. This enzyme
CC   is a member of the ABC transporter family.
XX
SQ   Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;

Alignment Scores:
Pred. No.:              1.61e-283     Length:         4279
Score:                  3112.50       Matches:        601
Percent Similarity:     75.02%        Conservative:   201
Best Local Similarity:  56.22%        Mismatches:     220
Query Match:            58.40%        Indels:         47
DB:                     22            Gaps:           6

US-09-873-409-4 (1-1058) x AAD03505 (1-4279)
```

```
Qy      4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
          |||   |||   ||||||||   |||||:::||||||||||||||||||||||||||||
Db    734 TCATTTACTGATAAAGAACTCTTGGCCTATGCAAAAGCTGGAGCAGTAGCTGAAGAAGTC 793

Qy     24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
          |||::::::||||||||||||||||||||||||       |||:::|||||||||:::
Db    794 TTAGCAGCAATCAGAACTGTGATTGCCTTTGGAGGACAAAAGAAAGAACTTGAA------ 847

Qy     44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63

Db    847 ------------------------------------------------------------ 847

Qy     64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
                         ||||||   :::||||||::::::||||||     ||||||||
Db    848 ---------------AGGTACAACAAAAATTTAGAAGAAGCTAAAGGAATTGGGATAAAG 892

Qy     84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
          :::   |||   :::   :::|||:::||||||   :::     :::     :::|||
Db    893 AAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCATCATAT 952

Qy    104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
             ||||||||||||||||||||||||||||:::|||:::   |||      |||:::|||
Db    953 GCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TATTCTATT 1006

Qy    124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
          |||   ||||||   |||||||||||||||:::      ::::::   ||||||   |||
Db   1007 GGACAAGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTAGTATTGGACAGGCA 1066

Qy    144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
             |||      |||   ||||||   |||||||||||||||:::   ||||||::::::
Db   1067 TCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1126

Qy    164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
          ||||||   ||||||||||||||||||:::::::::|||:::||||||:::::::|||
Db   1127 ATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGATAATATT 1186

Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
          :::|||   :::|||||||||||||||   |||:::|||||||||||||      :::|||
Db   1187 AAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAAGTTAAG 1246

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
          |||||||||||||||||||||||||:::::::::||||||:::|||||||||||||||||
Db   1247 ATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTTGGGAAC 1306

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
          :::|||   ||||||||||||   ||||||||||:::||||||||||||||||||   |||
Db   1307 AGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCCACAGAT 1366

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
          |||   :::   :::|||      |||||||||   :::||||||||||||   |||:::
Db   1367 GGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTTCGGGAA 1426

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
             |||||||||||||||||||||||||||||||   |||||||||:::   ||||||
Db   1427 ATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAAAACATT 1486

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
          :::|||||||||::::::||||||   :::|||:::|||:::|||   :::|||||||||
Db   1487 CGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAAT 1546

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
          ||||||||||||||||||:::   ||||||||||||:::||||||||||||||||:::|||
Db   1547 GCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAGAGAGGG 1606

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
          ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1607 GCCCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTTCGCAAC 1666

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
          ||||||||||:::||||||||||||||||||||||||||||:::||||||::::::
Db   1667 CCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTG 1726

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
```

```
Db   1727 GTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCTCATCGT 1786

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db   1787 TTGTCTACAGTTCGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATTGTGGAG 1846

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   1847 AAAGGAAATCATGATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTCACAATG 1906

Qy    424 GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 443
Db   1907 CAG---ACAAGAGGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAAAGTGAA 1963

Qy    444 ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--------- 460
Db   1964 AGTGATGCCTTGGAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAAAGAAGA 2020

Qy    461 -----------------------------------AspLysAlaGluGluSerThrGln 468
Db   2021 TCAACTCGCAGGAGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACAAAAGAG 2080

Qy    469 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 488
Db   2081 GACTTGAATGAGAATGTACCTCCAGTTTCCTTCTGGAGGATTCTGAAGCTGAACTCAACT 2140

Qy    489 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 508
Db   2141 GAATGGCCTTATTTTGTGGTTGGTATATTTTGTGCTATTATAAACGGAGGCCTGCAACCA 2200

Qy    509 ValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThr 527
Db   2201 GCATTTTCAATAATATTTTCAAGGATTATAGGGATCTTTACCCGAGATGAGGATCCTGAA 2260

Qy    528 ThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 547
Db   2261 ACAAAACGACAGAATAGTAACATGTTTTCTGTATTGTTTCTAGTCCTTGGAATTATTTCT 2320

Qy    548 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 567
Db   2321 TTTATTACATTTTTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTCACTAAG 2380

Qy    568 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 587
Db   2381 CGGCTTCGATACATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTTGATGAC 2440

Qy    588 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 607
Db   2441 CCTAAAAACACCACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAAGTTAAA 2500

Qy    608 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 627
Db   2501 GGGGCTATAGGTTCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGGACAGGC 2560

Qy    628 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 647
Db   2561 ATTATTATATCCTTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATTGTACCC 2620

Qy    648 ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp 667
Db   2621 ATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGAT 2680

Qy    668 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 687
Db   2681 AAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTCCGAACT 2740

Qy    688 IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 707
Db   2741 GTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTGCAAGTA 2800

Qy    708 GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis 727
Db   2801 CCATACAGAAACTCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATCACCCAG 2860

Qy    728 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla 747
Db   2861 GCAATGATGTATTTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTGGCAAAT 2920

Qy    748 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 767
Db   2921 GAGTTCATGAACTTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGTGCCATG 2980

Qy    768 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 787
Db   2981 GCAGTGGGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCAGCAGCC 3040

Qy    788 HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys 807
Db   3041 CACGTCATCATGATCATTGAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCACGGCCTC 3100

Qy    808 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys 827
Db   3101 AAGCCAAATACGTTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTATCCCACT 3160

Qy    828 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 847
Db   3161 CGACCAGACATCCCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAGACGCTG 3220

Qy    848 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 867
Db   3221 GCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAGCGCTTC 3280

Qy    868 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 887
Db   3281 TATGACCCCTTGGCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTGAATGTC 3340

Qy    888 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 907
Db   3341 CAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGACTGCAGC 3400

Qy    908 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 927
Db   3401 ATTGCCGAGAACATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAGATTATG 3460

Qy    928 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn 947
Db   3461 CAGGCAGCCAAGGAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAATACAAC 3520

Qy    948 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 967
Db   3521 ACCAGAGTAGGAGACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATTGCCATA 3580

Qy    968 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 987
Db   3581 GCTCGCGCTCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCTACATCAGCTCTG 3640

Qy    988 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1007
Db   3641 GATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACC 3700

Qy   1008 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1027
Db   3701 TGCATTGTGATCGCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTGGTGTTT 3760

Qy   1028 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1047
Db   3761 CAGAATGGCAAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGGCCCAGAAAGGCATC 3820

Qy   1048 TyrPheLysLeuValAsnAlaGlnSer 1056
Db   3821 TATTTTTCCATGGTCAGTGTCCAGGCT 3847
```

RESULT 6
AAD03489
ID AAD03489 standard; cDNA; 4317 BP.

```
XX
AC   AAD03489;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) cDNA #2.
XX
KW   Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             70..3912
FT                   /*tag=  a
FT                   /product= "Dog P-glycoprotein (PGP) #2"
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00304.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 1; Page 66-72; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) cDNA. The
CC   PGP enzyme functions as an efflux pump exporting small molecules
CC   across the cell membrane. This enzyme is a member of the ABC
CC   transporter family.
XX
SQ   Sequence 4317 BP; 1293 A; 844 C; 1019 G; 1161 T; 0 other;

Alignment Scores:
Pred. No.:              2.03e-283      Length:         4317
Score:                  3111.50        Matches:        600
Percent Similarity:     75.02%         Conservative:   202
Best Local Similarity:  56.13%         Mismatches:     220
Query Match:            58.38%         Indels:         47
DB:                     22             Gaps:           6

US-09-873-409-4 (1-1058) x AAD03489 (1-4317)

Qy      4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
Db    784 TCATTTACTGATAAAGAACTCTTGGCCTATGCAAAAGCTGGAGCAGTAGCTGAAGAAGTC 843

Qy     24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43

Db    844 TTAGCAGCAATCAGAACTGTGATTGCCTTTGGAGGACAAAAGAAAGAACTTGAA------ 897

Qy     44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63

Db    897 ------------------------------------------------------------ 897

Qy     64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83

Db    898 --------------AGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 942

Qy     84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103

Db    943 AAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCATCATAT 1002

Qy    104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123

Db   1003 GCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TATACTATT 1056

Qy    124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143

Db   1057 GGACAGGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGACAGGCA 1116

Qy    144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163

Db   1117 TCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1176

Qy    164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183

Db   1177 ATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGATAATATT 1236

Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203

Db   1237 AAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAAGTTAAG 1296

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223

Db   1297 ATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTTGGGAAC 1356

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243

Db   1357 AGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCCACAGAT 1416

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263

Db   1417 GGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTTCGGGAA 1476

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283

Db   1477 ATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAAAACATT 1536

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303

Db   1537 CGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAAT 1596

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323

Db   1597 GCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAGAGAGGG 1656

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343

Db   1657 GCCCGGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTTCGCAAC 1716

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363

Db   1717 CCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTG 1776

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383

Db   1777 GTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCTCATCGT 1836

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
```

```
Db   1837 TTGTCTACAGTTCGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATTGTGGAG 1896

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   1897 AAAGGAAATCATGATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTCACAATG 1956

Qy    424 GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 443
Db   1957 CAG---ACAAGAGGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAAAGTGAA 2013

Qy    444 ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--------- 460
Db   2014 AGTGATGCCTTGGAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAAAGAAGA 2070

Qy    461 ------------------------------------AspLysAlaGluGluSerThrGln 468
Db   2071 TCAACTCGCAGGAGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACAAAAGAG 2130

Qy    469 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 488
Db   2131 GACTTGAATGAGAATGTTCCTTCAGTTTCCTTCTGGAGGATTCTGAAGCTGAACTCAACT 2190

Qy    489 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 508
Db   2191 GAATGGCCTTATTTGTGGTTGGTATATTTGTGCTATTATAAACGGAGGCCTGCAACCA 2250

Qy    509 ValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThr 527
Db   2251 GCATTTTCAATAATATTTTCAAGGATTATAGGGATCTTTACCCGAGATGAGGATCCTGAA 2310

Qy    528 ThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 547
Db   2311 ACAAAACGACAGAATAGTAACATGTTTTCTGTATTGTTTCTAGTCCTTGGAATTATTTCT 2370

Qy    548 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 567
Db   2371 TTTATTACATTTTTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTCACTAAG 2430

Qy    568 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 587
Db   2431 CGGCTTCGATACATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTTGATGAC 2490

Qy    588 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 607
Db   2491 CCTAAAAACACCACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAAGTTAAA 2550

Qy    608 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 627
Db   2551 GGGGCTATAGGTTCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGGACAGGC 2610

Qy    628 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 647
Db   2611 ATTATTATATCCTTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATTGTACCC 2670

Qy    648 ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp 667
Db   2671 ATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGAT 2730

Qy    668 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 687
Db   2731 AAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTCCGAACT 2790

Qy    688 IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 707
Db   2791 GTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTGCAAGTA 2850

Qy    708 GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis 727
Db   2851 CCATACAGAAACTCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATCACCCAG 2910

Qy    728 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla 747
Db   2911 GCAATGATGTATTTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTGGCAAAT 2970

Qy    748 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 767
Db   2971 GAGTTCATGAACTTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGTGCCATG 3030

Qy    768 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 787
Db   3031 GCAGTGGGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCAGCAGCC 3090

Qy    788 HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys 807
Db   3091 CACGTCATCATGATCATTGAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCACGGCCTC 3150

Qy    808 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys 827
Db   3151 AAGCCAAATACGTTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTATCCCACT 3210

Qy    828 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 847
Db   3211 CGACCAGACATCCCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAGACGCTG 3270

Qy    848 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 867
Db   3271 GCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAGCGCTTC 3330

Qy    868 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 887
Db   3331 TATGACCCCTTGGCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTGAATGTC 3390

Qy    888 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 907
Db   3391 CAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGACTGCAGC 3450

Qy    908 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 927
Db   3451 ATTGCCGAGAACATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAGATTATG 3510

Qy    928 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn 947
Db   3511 CAGGCAGCCAAGGAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAATACAAC 3570

Qy    948 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 967
Db   3571 ACCAGAGTAGGAGACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATTGCCATA 3630

Qy    968 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 987
Db   3631 GCTCGCGCTCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCTACATCAGCTCTG 3690

Qy    988 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1007
Db   3691 GATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACC 3750

Qy   1008 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1027
Db   3751 TGCATTGTGATCGCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTGGTGTTT 3810

Qy   1028 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1047
Db   3811 CAGAATGGCAAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGGCTCAGAAAGGCATC 3870

Qy   1048 TyrPheLysLeuValAsnAlaGlnSer 1056
Db   3871 TATTTTTCCATGATCAGTGTCCAGGCT 3897
```

RESULT 7
AAZ52047

```
ID   AAZ52047 standard; cDNA; 4369 BP.
XX
AC   AAZ52047;
XX
DT   18-JUL-2000  (first entry)
XX
```

```
DE   Rat multidrug resistance protein 1a cDNA.
XX
KW   Multidrug resistance protein 1a; mdr1a; multi-specific drug transporter;
KW   drug formulation; formulation excipient; compound design; inflammation;
KW   pharmacokinetic; oral absorption; cancer; cardiovascular disease;
KW   central nervous system disorder; auto-immune disease; kidney disease; ss.
XX
OS   Rattus rattus.
XX
FH   Key             Location/Qualifiers
FT   CDS             352..4170
FT                   /*tag=  a
FT                   /product= "Rat multidrug resistance protein"
XX
PN   WO200015650-A1.
XX
PD   23-MAR-2000.
XX
PF   10-SEP-1999;   99WO-US20770.
XX
PR   17-SEP-1998;   98US-0156800.
PR   09-DEC-1998;   98US-0208809.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI   Brun KA,  Chenery RJ,  Ellens H,  Feild JA,  Yue L;
XX
DR   WPI; 2000-271372/23.
DR   P-PSDB; AAY70596.
XX
PT   Isolated rat mdr1a polynucleotides and polypeptides, useful in assays
PT   to provide information on drug formulation, selection of formulation
PT   excipients and compound design -
XX
PS   Claim 2; Page 23-24; 33pp; English.
XX
CC   The present cDNA sequence encodes rat multidrug resistance
CC   protein 1a (mdr1a). This protein is a member of the multi-specific drug
CC   transporters family. Mdr1a is used in assays to provide information on
CC   drug formulation, selection of formulation excipients and compound
CC   design. They are used in cell based, membrane based, binding or other
CC   assays to provide information that may enhance drug formulation. This
CC   invention further relates to the generation of in vivo and in vitro
CC   comparison data to predict oral absorption and pharmacokinetics. This
CC   enables the selection of drugs with optimal pharmacokinetics, i.e. good
CC   oral bioavailability, brain penetration, plasma half life, and minimum
CC   drug interaction. Transgenic and knock-out animals created using rat
CC   mdr1a provides an insight into treating and preventing human diseases
CC   including cancer, inflammation, cardiovascular disease, central nervous
CC   system disorders, auto-immune and kidney disease.
XX
SQ   Sequence 4369 BP; 1241 A; 973 C; 1096 G; 1059 T; 0 other;

Alignment Scores:
Pred. No.:              6.13e-283      Length:          4369
Score:                  3106.50        Matches:         600
Percent Similarity:     75.54%         Conservative:    206
Best Local Similarity:  56.23%         Mismatches:      218
Query Match:            58.28%         Indels:          43
DB:                     21             Gaps:            7

US-09-873-409-4 (1-1058) x AAZ52047 (1-4369)

Qy       4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
Db    1039 TCATTTACTGATAAGGAACTCCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTC 1098

Qy      24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db    1099 TTAGCAGCCATCAGAACTGTGATTGCCTTTGGAGGACAAAAGAAGGAACTTGAA------ 1152

Qy      44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db    1152 ------------------------------------------------------------ 1152

Qy      64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db    1153 ---------------AGGTACAATAACAATTTGGAAGAAGCTAAAAGGCTTGGGATAAAG 1197

Qy      84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db    1198 AAAGCTATCACGGCCAACATTTCCATGGGTGCAGCTTTTCTGCTTATCTATGCATCATAT 1257

Qy     104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db    1258 GCTCTGGCATTCTGGTATGGGACTTCCTTGGTCATCTCAAAAGAA------TACACTATT 1311

Qy     124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db    1312 GGACAAGTGCTCACTGTCTTTTTTTCTGTATTAATTGGAGCATTCAGTGTTGGGCAGGCA 1371

Qy     144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db    1372 TCTCCAAATATTGAAGCCTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTCAGTATA 1431

Qy     164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db    1432 ATTGATAATAAGCCCAGTATAGACAGCTTCTCAAAGAGTGGGCACAAACCCGACAACATA 1491

Qy     184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db    1492 CAAGGAAATTTGGAATTCAAAAATATTCACTTCAGTTACCCGTCTCGAAAAGACGTTCAG 1551

Qy     204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db    1552 ATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTGGTTGGCAAC 1611

Qy     224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db    1612 AGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCATAGAG 1671

Qy     244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db    1672 GGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAA 1731

Qy     264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db    1732 ATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCCACCACAATTGCCGAAAACATT 1791

Qy     284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db    1792 CGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAAT 1851

Qy     304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db    1852 GCCTATGATTTCATCATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAGAGAGGG 1911

Qy     324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db    1912 GCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAAC 1971

Qy     344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db    1972 CCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTG 2031

Qy     364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db    2032 GTTCAGGCCGCTCTGGATAAGGCTAGAGAAGGCCGGACCACCATTGTGATAGCTCACCGC 2091

Qy     384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db    2092 TTGTCTACAGTTCGCAATGCTGACGTCATTGCTGGTTTTGATGGTGGTGTCATTGTGGAG 2151

Qy     404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db    2152 CAAGGAAATCATGATGAGCTCATGAGAGAGAAAGGAATTTACTTCAAACTTGTCATGACT 2211
```

```
Qy    424 Gln------------AspIleLysLysAlaAspGluGlnMetGluSer----------- 435
Db   2212 CAGACAGCAGGAAATGAAATTGAATTAGGAAATGAAGCTTGTGAATCTAAAGACGGAATT 2271

Qy    436 ------MetThrTyrSerThrGluArgLysThrAsnSerLeu------ProLeuHisSer 451
Db   2272 GATAATGTGGACATGTCTTCAAAAGATTCGGGATCCAGTCTAATAAGAAGAAGATCAACT 2331

Qy    452 ValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLys 470
Db   2332 CGCAAAAGCATCCGTGGGCCACATGATCAAGACGGGGAACTTAGCACCAAAGAGGCTCTG 2391

Qy    471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
Db   2392 GATGACGACGTACCTCCAGCTTCCTTTTGGCGGATCCTGAAGTTGAATTCAACTGAATGG 2451

Qy    491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db   2452 CCTTATTTTGTGGTTGGTGTATTTTGTGCCATAATAAATGGAGGCTTGCAACCAGCATTC 2511

Qy    511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeu--- 529
Db   2512 TCCATAATATTTTCAAAGGTTGTAGGGGTTTTTACAAAAAATGACACCCCTGAAATCCAG 2571

Qy    530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db   2572 CGGCAGAACAGCAACTTGTTTCTTTATTGTTCCTGATCCTTGGGATCATCTCTTTCATT 2631

Qy    550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db   2632 ACGTTTTTCCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGACTC 2691

Qy    570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
Db   2692 CGATACATGGTCTTCAAATCCATGCTGAGACAGGACATAAGCTGGTTTGATGACCCTAAA 2751

Qy    590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db   2752 AACACCACAGGAGCGCTGACCACCAGGCTTGCCAATGACGCTGCTCAAGTGAAAGGGGCT 2811

Qy    610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db   2812 ACAGGGTCTAGGCTTGCTGTTATTACCCAGAACATAGCAAATCTTGGGACAGGCATCATC 2871

Qy    630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db   2872 ATATCCCTGATCTACGGCTGGCAATTGACACTTTTACTCCTAGCAATTGTTCCCATCATT 2931

Qy    650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db   2932 GCTATAGCAGGAGTGGTTGAAATGAAAATGTTGTCTGGACAAGCGCTGAAAGATAAGAAG 2991

Qy    670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db   2992 GAACTAGAAGGTTCTGGGAAGATCGCTACAGAAGCAATTGAAAACTTTCGCACTGTCGTC 3051

Qy    690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db   3052 TCTTTGACTCGGGAGCAGAAGTTTGAAACTATGTATGCCCAGAGCTTGCAGATACCATAC 3111

Qy    710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db   3112 AGAAATGCTTTGAAGAAAGCGCACGTCTTTGGGATCACTTTCTCCTTCACCCAGGCCATG 3171

Qy    730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db   3172 ATGTATTTCTCCTATGCTGCTTGTTTCCGGTTTGATGCCTACTTGGTGGCACGAGAACTC 3231

Qy    750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db   3232 ATGACATTTGAAAATGTTCTGTTAGTATTCTCAGCTATTGTCTTTGGTGCCATGGCAGTG 3291

Qy    770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db   3292 GGGCAGGTCAGTTCATTCGCTCCTGACTACGCGAAAGCCAAAGTCTCGGCATCCCACATC 3351

Qy    790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db   3352 ATCAGGATCATTGAGAAAATCCCTGAGATTGACAGCTACAGCACGGAGGGCTTGAAGCCT 3411

Qy    810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db   3412 AATATGTTGGAAGGAAATGTGAAATTTAATGGAGTCATGTTCAACTATCCCACCCGACCC 3471

Qy    830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db   3472 AACATCCCAGTGCTTCAGGGGCTGAGCCTAGAGGTGAAGAAAGGGCAGACGCTGGCCCTC 3531

Qy    850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db   3532 GTGGGCAGCAGTGGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTTGAGCGCTTCTATGAC 3591

Qy    870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
Db   3592 CCCATGGCCGGAACAGTGTTTCTAGATGGCAAAGAAATAAAGCAACTCAATGTCCAGTGG 3651

Qy    890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db   3652 CTCCGCGCCCACCTGGGCATTGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATCGCC 3711

Qy    910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db   3712 GAGAACATTGCCTACGGAGACAACAGCCGTGTCGTGTCTCATAAGGAGATCGTGAAGGCA 3771

Qy    930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db   3772 GCCAAGGAGGCCAACATCCACCAGTTCATCGACTCACTGCCTGAGAAATACAACACCAGA 3831

Qy    950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db   3832 GTGGGAGACAAAGGGACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATCGCCATCGCGCGC 3891

Qy    970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db   3892 GCCCTCGTCAGACAGCCTCACATCTTACTTCTGGATGAAGCGACATCAGCTCTGGATACG 3951

Qy    990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db   3952 GAGAGTGAAAAGGTCGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACCTGCATT 4011

Qy   1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db   4012 GTGATCGCGCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATTCAGAAC 4071

Qy   1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
Db   4072 GGCCAGGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCCCAGAAAGGCATCTATTTC 4131

Qy   1050 LysLeuValAsnAlaGlnSer 1056
Db   4132 TCGATGGTCAGTGTGCAGGCT 4152
```

RESULT 8
AAZ52048

```
ID   AAZ52048 standard; cDNA; 4425 BP.
XX
AC   AAZ52048;
XX
DT   18-JUL-2000  (first entry)
XX
DE   Rat multidrug resistance protein 1a cDNA derived from EST sequences.
XX
KW   Multidrug resistance protein 1a; mdr1a; multi-specific drug transporter;
KW   drug formulation; formulation excipient; compound design; inflammation;
KW   pharmacokinetic; oral absorption; cancer; cardiovascular disease;
```

```
KW   central nervous system disorder; auto-immune disease; kidney disease;
KW   EST; expressed sequence tag; ss.
XX
OS   Rattus rattus.
XX
FH   Key             Location/Qualifiers
FT   CDS             352..4170
FT                   /*tag=  a
FT                   /product= "Rat multidrug resistance protein"
XX
PN   WO200015650-A1.
XX
PD   23-MAR-2000.
XX
PF   10-SEP-1999;    99WO-US20770.
XX
PR   17-SEP-1998;    98US-0156800.
PR   09-DEC-1998;    98US-0208809.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI   Brun KA,  Chenery RJ,  Ellens H,  Feild JA,  Yue L;
XX
DR   WPI; 2000-271372/23.
DR   P-PSDB; AAY70597.
XX
PT   Isolated rat mdr1a polynucleotides and polypeptides, useful in assays
PT   to provide information on drug formulation, selection of formulation
PT   excipients and compound design -
XX
PS   Claim 11; Page 27-29; 33pp; English.
XX
CC   The present cDNA sequence encodes rat multidrug resistance protein 1a
CC   (mdr1a). This cDNA is derived from EST (expressed sequence tag)
CC   sequences. Mdr1a is used in assays to provide information on
CC   drug formulation, selection of formulation excipients and compound
CC   design. They are used in cell based, membrane based, binding or other
CC   assays to provide information that may enhance drug formulation. This
CC   invention further relates to the generation of in vivo and in vitro
CC   comparision data to predict oral absorption and pharmacokinetics. This
CC   enables the selection of drugs with optimal pharmacokinetics, i.e. good
CC   oral bioavailability, brain penetration, plasma half life, and minimum
CC   drug interaction. Transgenic and knock-out animals created using rat
CC   mdr1a provides an insight into treating and preventing human diseases
CC   including cancer, inflammation, cardiovascular disease, central nervous
CC   system disorders, auto-immune and kidney disease.
XX
SQ   Sequence 4425 BP; 1250 A; 992 C; 1113 G; 1067 T; 2 U; 1 other;

Alignment Scores:
Pred. No.:              6.25e-283      Length:        4425
Score:                  3106.50        Matches:       600
Percent Similarity:     75.54%         Conservative:  206
Best Local Similarity:  56.23%         Mismatches:    218
Query Match:            58.28%         Indels:        43
DB:                     21             Gaps:          7

US-09-873-409-4 (1-1058) x AAZ52048 (1-4425)

Qy      4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
Db   1039 TCATTTACTGATAAGGAACTCCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTC 1098

Qy     24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db   1099 TTAGCAGCCATCAGAACTGTGATTGCCTTTGGAGGACAAAAGAAGGAACTTGAA------ 1152

Qy     44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db   1152 ------------------------------------------------------------ 1152

Qy     64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db   1153 ---------------AGGTACAATAACAATTTGGAAGAAGCTAAAAGGCTTGGGATAAAG 1197

Qy     84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db   1198 AAAGCTATCACGGCCAACATTTCCATGGGTGCAGCTTTTCTGCTTATCTATGCATCATAT 1257

Qy    104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db   1258 GCTCTGGCATTCTGGTATGGGACTTCCTTGGTCATCTCAAAAGAA------TACACTATT 1311

Qy    124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db   1312 GGACAAGTGCTCACTGTCTTTTTTTCTGTATTAATTGGAGCATTCAGTGTTGGGCAGGCA 1371

Qy    144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db   1372 TCTCCAAATATTGAAGCCTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTCAGTATA 1431

Qy    164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db   1432 ATTGATAATAAGCCCAGTATAGACAGCTTCTCAAAGAGTGGGCACAAACCCGACAACATA 1491

Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db   1492 CAAGGAAATTTGGAATTCAAAAATATTCACTTCAGTTACCCGTCTCGAAAAGACGTTCAG 1551

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db   1552 ATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTGGTTGGCAAC 1611

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db   1612 AGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCATAGAG 1671

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db   1672 GGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAA 1731

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db   1732 ATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCCACCACAATTGCCGAAAACATT 1791

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db   1792 CGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAAT 1851

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db   1852 GCCTATGATTTCATCATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAGAGAGGG 1911

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db   1912 GCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAAC 1971

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db   1972 CCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTG 2031

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db   2032 GTTCAGGCCGCTCTGGATAAGGCTAGAGAAGGCCGGACCACCATTGTGATAGCTCACCGC 2091

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db   2092 TTGTCTACAGTTCGCAATGCTGAUGTCATTGCTGGTTTTGATGGTGGTGTCATTGTGGAG 2151

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   2152 CAAGGAAATCATGATGAGCTCATGAGAGAGAAAGGAATTTACTTCAAACTTGTCATGACT 2211

Qy    424 Gln-----------AspIleLysLysAlaAspGluGlnMetGluSer------------ 435
Db   2212 CAGACAGCAGGAAATGAAATTGAATTAGGAAATGAAGCTTGTGAATCTAAAGAUGGAATT 2271
```

```
Qy     436 ------MetThrTyrSerThrGluArgLysThrAsnSerLeu------ProLeuHisSer 451
                 :::      |||::::::        :::|||||              :::
Db    2272 GATAATGTGGACATGTCTTCAAAAGATTCRGGATCCAGTCTAATAAGAAGAAGATCAACT 2331

Qy     452 ValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLys 470
              |||||||||:::                     |||   ||||||   ::::::
Db    2332 CGCAAAAGCATCCGTGGGCCACATGATCAAGACGGGGAACTTAGCACCAAAGAGGCTCTG 2391

Qy     471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
           :::       :::|||      |||      :::||||||||||||||       ||||||
Db    2392 GATGACGACGTACCTCCAGCTTCCTTTTGGCGGATCCTGAAGTTGAATTCAACTGAATGG 2451

Qy     491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
           |||:::    |||:::|||                :::   |||   |||
Db    2452 CCTTATTTTGTGGTTGGTGTATTTTGTGCCATAATAAATGGAGGCTTGCAACCAGCATTC 2511

Qy     511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeu--- 529
           |||||||||||:::|||::::::    :::|||      ||||||         :::
Db    2512 TCCATAATATTTCAAAGGTTGTAGGGGTTTTTACAAAAAATGACACCCCTGAAATCCAG 2571

Qy     530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
           :::   ::::::   ::::::|||::::::|||:::||||||||:::|||   |||:::
Db    2572 CGGCAGAACAGCAACTTGTTTTCTTTATTGTTCCTGATCCTTGGGATCATCTCTTTCATT 2631

Qy     550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
           ::::::|||:::||||||      :::|||:::||||||||||||||||||   ||||||
Db    2632 ACGTTTTTCCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGACTC 2691

Qy     570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
           |||::::::   ||||||:::||||||    |||||||||:::||||||:::   :::
Db    2692 CGATACATGGTCTTCAAATCCATGCTGAGACAGGACATAAGCTGGTTTGATGACCCTAAA 2751

Qy     590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
           |||:::||||||    ||||||       |||   |||:::::::::||||||
Db    2752 AACACCACAGGAGCGCTGACCACCAGGCTTGCCAATGACGCTGCTCAAGTGAAAGGGGCT 2811

Qy     610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
           ||||||||||||:::    |||:::|||||||||      |||:::|||      :::|||
Db    2812 ACAGGGTCTAGGCTTGCTGTTATTACCCAGAACATAGCAAATCTTGGGACAGGCATCATC 2871

Qy     630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
           ||||||   ||||||||||||:::::::|||   |||:::|||:::|||   |||::::::
Db    2872 ATATCCCTGATCTACGGCTGGCAATTGACACTTTTACTCCTAGCAATTGTTCCCATCATT 2931

Qy     650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
           |||:::   |||::::::|||            ::::::|||   |||   ||||||:::
Db    2932 GCTATAGCAGGAGTGGTTGAAATGAAAATGTTGTCTGGACAAGCGCTGAAAGATAAGAAG 2991

Qy     670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
           ||||||:::      :::||||||||||||||||||:::||||||   ||||||:::|||
Db    2992 GAACTAGAAGGTTCTGGGAAGATCGCTACAGAAGCAATTGAAAACTTTCGCACTGTCGTC 3051

Qy     690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
           ||||||||||||||:::   ||||||   ||||||   :::   ||||||      :::
Db    3052 TCTTTGACTCGGGAGCAGAAGTTTGAAACTATGTATGCCCAGAGCTTGCAGATACCATAC 3111

Qy     710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
           ||||||      ||||||||:::   :::   |||      ::::::|||:::   |||
Db    3112 AGAAATGCTTTGAAGAAAGCGCACGTCTTTGGGATCACTTTCTCCTTCACCCAGGCCATG 3171

Qy     730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
           :::||||||:::|||||||||   |||||||||   |||||||||:::
Db    3172 ATGTATTTCTCCTATGCTGCTTGTTTCCGGTTTGATGCCTACTTGGTGGCACGAGAACTC 3231

Qy     750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
           ||||||   |||      :::   :::|||||||||:::||||||   :::||||||||||:::
Db    3232 ATGACATTTGAAAATGTTCTGTTAGTATTCTCAGCTATTGTCTTTGGTGCCATGGCAGTG 3291

Qy     770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
           |||:::              |||||:::|||:::|||||||||      |||:::|||:::
Db    3292 GGGCAGGTCAGTTCATTCGCTCCTGACTACGCGAAAGCCAAAGTCTCGGCATCCCACATC 3351

Qy     790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
                 ::::::||||||   |||   |||||||||   |||   ||||||   ||||||
Db    3352 ATCAGGATCATTGAGAAAATCCCTGAGATTGACAGCTACAGCACGGAGGGCTTGAAGCCT 3411

Qy     810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
           :::      |||||||||::::::|||   |||   |||   ||||||   ||||||
Db    3412 AATATGTTGGAAGGAAATGTGAAATTTAATGGAGTCATGTTCAACTATCCCACCCGACCC 3471

Qy     830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
           ::::::   :::|||:::|||||||||       :::::::::|||:::|||:::|||
Db    3472 AACATCCCAGTGCTTCAGGGGCTGAGCCTAGAGGTGAAGAAGGGCAGACGCTGGCCCTC 3531

Qy     850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
           |||||||||||||||||||||||||||       ||||||||||||:::|||   ||||||
Db    3532 GTGGGCAGCAGTGGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTTGAGCGCTTCTATGAC 3591

Qy     870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
           |||:::   |||   |||      ||||||   :::   |||:::||||||||||||||||
Db    3592 CCCATGGCCGGAACAGTGTTTCTAGATGGCAAAGAAATAAAGCAACTCAATGTCCAGTGG 3651

Qy     890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
           ||||||:::   :::   ||||||   ||||||||||:::||||||:::||||||||||||
Db    3652 CTCCGCGCCCACCTGGGCATTGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATCGCC 3711

Qy     910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
           |||||||||||||||||||||||||||||||||||||         ||||||   :::|||
Db    3712 GAGAACATTGCCTACGGAGACAACAGCCGTGTCGTGTCTCATAAGGAGATCGTGAAGGCA 3771

Qy     930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
           |||      ||||||||||||   ||||||:::   ||||||||||||||||||||:::
Db    3772 GCCAAGGAGGCCAACATCCACCAGTTCATCGACTCACTGCCTGAGAAATACAACACCAGA 3831

Qy     950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
           ||||||   ||||||   ||||||||||||||||||||||||||:::||||||||||||
Db    3832 GTGGGAGACAAAGGGACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATCGCCATCGCGCGC 3891

Qy     970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
           ||||||::::::::|||   |||||||||||||||||||||||||||||||||
Db    3892 GCCCTCGTCAGACAGCCTCACATCTTACTTCTGGATGAAGCGACATCAGCTCTGGATACG 3951

Qy     990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
           :::||||||||||||||||||   ||||||||||||||||||   ||||||||||||:::
Db    3952 GAGAGTGAAAAGGTCGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACCTGCATT 4011

Qy    1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
           |||:::   ||||||||||||   ||||||||||||||||||||||||||||:::   |||
Db    4012 GTGATCGCGCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATTCAGAAC 4071

Qy    1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
           |||::::::||||||   ||||||||||||:::||||||      :::   |||||||||
Db    4072 GGCCAGGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCCCAGAAAGGCATCTATTTC 4131

Qy    1050 LysLeuValAsnAlaGlnSer 1056
              :::|||:::   |||:::
Db    4132 TCGATGGTCAGTGTGCAGGCT 4152
```

```
RESULT 9
AAZ49332
ID   AAZ49332 standard; cDNA; 3860 BP.
XX
AC   AAZ49332;
XX
DT   14-MAR-2000  (first entry)
XX
DE   Human wild-type multidrug resistance-1 (MDR-1) cDNA.
XX
KW   Multidrug resistance; MDR-1; P-glycoprotein;
KW   transmembrane efflux pump; haematopoietic stem cell; transduction;
KW   bone marrow transplantation; chemotherapy; radiation therapy; cancer;
KW   gene therapy; gene replacement; genetic defect; thalassaemia;
KW   Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
KW   cytokine; wild-type; ds.
XX
```

```
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3843
FT                   /*tag= a
FT                   /product= "Human wild-type MDR-1 protein"
FT   mutation        replace (553..555, GTT)
FT                   /*tag= b
FT                   /note= "cDNA sequence of G185V human mutant MDR-1 given
FT                          in AAZ49333"
XX
PN   WO9961589-A2.
XX
PD   02-DEC-1999.
XX
PF   27-MAY-1999;   99WO-US11825.
XX
PR   28-MAY-1998;   98US-0086988.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI   Sorrentino B,  Bunting K;
XX
DR   WPI; 2000-072615/06.
DR   P-PSDB; AAY58186.
XX
PT   Ex vivo expansion of hematopoietic stem cells transduced with a
PT   sequence encoding human multidrug resistance-1, used for bone marrow
PT   transplantation -
XX
PS   Claim 10; Page 68-70; 113pp; English.
XX
CC   This sequence represents cDNA encoding human wild-type
CC   multidrug resistance protein MDR-1. MDR-1 is a transmembrane
CC   efflux pump, responsible for the export of drugs from cells,
CC   particularly cancer cells. Wild-type MDR-1 shows increased
CC   resistance to etoposide and decreased resistance to vinca
CC   alkaloids compared with a mutant form (AAY58187) where the Gly at
CC   position 185 is replaced by Val. The invention relates to transducing
CC   haematopoietic stem cells with nucleic acid encoding an MDR protein
CC   and culturing the modified cells. The modified haematopoietic stem
CC   cells are useful in bone marrow transplantation (to reconstitute
CC   haematopoietic systems in patients who have undergone chemotherapy or
CC   radiation therapy) and in ex vivo gene therapy of genetic defects in
CC   cells derived from haematopoietic stem cells, e.g., thalassaemia,
CC   Gauchier's disease, sickle cell anaemia or leukaemia. The modified
CC   cells can also be used to identify factors involved in regulating
CC   proliferation and differentiation in haematopoietic stem cells.
CC   Haematopoietic stem cells that express MDR-1 will be protected against
CC   chemotherapeutic agents, so can be engrafted while the patient is
CC   undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells
CC   provides sufficient cells to permit standard biochemical analysis.
CC   Overexpression of MDR-1 allows cytokine-driven expansion of
CC   haematopoietic stem cells by at least 10-fold compared with a maximum
CC   of 4-fold in known procedures.
XX
SQ   Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;

Alignment Scores:
Pred. No.:              1.88e-282     Length:         3860
Score:                  3100.50       Matches:        595
Percent Similarity:     75.54%        Conservative:   211
Best Local Similarity:  55.76%        Mismatches:     218
Query Match:            58.17%        Indels:         43
DB:                     21            Gaps:           6

US-09-873-409-4 (1-1058) x AAZ49332 (1-3860)

Qy      4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
Db    712 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 771

Qy     24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db    772 TTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAAAAGAAAGAACTTGAA------ 825

Qy     44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db    825 ------------------------------------------------------------ 825

Qy     64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db    826 ---------------AGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 870

Qy     84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db    871 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 930

Qy    104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db    931 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 984

Qy    124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db    985 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1044

Qy    144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db   1045 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104

Qy    164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db   1105 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1164

Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db   1165 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1224

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db   1225 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1284

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db   1285 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1344

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db   1345 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1404

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db   1405 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1464

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db   1465 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1524

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db   1525 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1584

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db   1585 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1644

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db   1645 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1704

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db   1705 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 1764

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
```

```
Db   1765 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 1824

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   1825 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 1884

Qy    424 Gln------------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
Db   1885 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 1944

Qy    438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 452
Db   1945 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2004

Qy    453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470
Db   2005 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2064

Qy    471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
Db   2065 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2124

Qy    491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db   2125 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2184

Qy    511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
Db   2185 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2244

Qy    530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db   2245 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2304

Qy    550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db   2305 ACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2364

Qy    570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
Db   2365 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2424

Qy    590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db   2425 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2484

Qy    610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db   2485 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2544

Qy    630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db   2545 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 2604

Qy    650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db   2605 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 2664

Qy    670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db   2665 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 2724

Qy    690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db   2725 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2784

Qy    710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db   2785 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 2844

Qy    730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db   2845 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 2904

Qy    750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db   2905 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 2964

Qy    770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db   2965 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3024

Qy    790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db   3025 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3084

Qy    810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db   3085 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3144

Qy    830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db   3145 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3204

Qy    850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db   3205 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3264

Qy    870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
Db   3265 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3324

Qy    890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db   3325 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3384

Qy    910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db   3385 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3444

Qy    930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db   3445 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3504

Qy    950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db   3505 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3564

Qy    970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db   3565 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 3624

Qy    990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db   3625 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 3684

Qy   1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db   3685 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 3744

Qy   1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
Db   3745 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 3804

Qy   1050 LysLeuValAsnAlaGlnSer 1056
Db   3805 TCAATGGTCAGTGTCCAGGCT 3825
```

RESULT 10
AAZ49333
ID AAZ49333 standard; cDNA; 3860 BP.
XX
AC AAZ49333;
XX
DT 14-MAR-2000 (first entry)
XX

```
DE   Human G185V mutant multidrug resistance-1 (MDR-1) cDNA.
XX
KW   Multidrug resistance; MDR-1; P-glycoprotein;
KW   transmembrane efflux pump; haematopoietic stem cell; transduction;
KW   bone marrow transplantation; chemotherapy; radiation therapy; cancer;
KW   gene therapy; gene replacement; genetic defect; thalassaemia;
KW   Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
KW   cytokine; mutant; ds.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3843
FT                   /*tag=  a
FT                   /product= "Human G185V mutant MDR-1 protein"
XX
PN   WO9961589-A2.
XX
PD   02-DEC-1999.
XX
PF   27-MAY-1999;   99WO-US11825.
XX
PR   28-MAY-1998;   98US-0086988.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI   Sorrentino B,  Bunting K;
XX
DR   WPI; 2000-072615/06.
DR   P-PSDB; AAY58187.
XX
PT   Ex vivo expansion of hematopoietic stem cells transduced with a
PT   sequence encoding human multidrug resistance-1, used for bone marrow
PT   transplantation -
XX
PS   Example 1; Page 79-82; 113pp; English.
XX
CC   This sequence represents cDNA encoding human G185V mutant multidrug
CC   resistance protein MDR-1, where the the Gly residue at postion 185
CC   of the wild-type protein (AAY58186) is replaced by Val. MDR-1 is a
CC   transmembrane efflux pump, responsible for the export of drugs from
CC   cells, particularly cancer cells. The wild-type MDR-1 shows increased
CC   resistance to etoposide and decreased resistance to vinca alkaloids
CC   compared with the G185V mutant. The invention relates to transducing
CC   haematopoietic stem cells with nucleic acid encoding an MDR protein
CC   and culturing the modified cells. The modified haematopoietic stem
CC   cells are useful in bone marrow transplantation (to reconstitute
CC   haematopoietic systems in patients who have undergone chemotherapy or
CC   radiation therapy) and in ex vivo gene therapy of genetic defects in
CC   cells derived from haematopoietic stem cells, e.g., thalassaemia,
CC   Gauchier's disease, sickle cell anaemia or leukaemia. The modified
CC   cells can also be used to identify factors involved in regulating
CC   proliferation and differentiation in haematopoietic stem cells.
CC   Haematopoietic stem cells that express MDR-1 will be protected against
CC   chemotherapeutic agents, so can be engrafted while the patient is
CC   undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells
CC   provides sufficient cells to permit standard biochemical analysis.
CC   Overexpression of MDR-1 allows cytokine-driven expansion of
CC   haematopoietic stem cells by at least 10-fold compared with a maximum
CC   of 4-fold in known procedures.
XX
SQ   Sequence 3860 BP; 1135 A; 746 C; 956 G; 1023 T; 0 other;

Alignment Scores:
Pred. No.:              1.88e-282      Length:         3860
Score:                  3100.50        Matches:        595
Percent Similarity:     75.54%         Conservative:   211
Best Local Similarity:  55.76%         Mismatches:     218
Query Match:            58.17%         Indels:         43
DB:                     21             Gaps:           6

US-09-873-409-4 (1-1058) x AAZ49333 (1-3860)
```

```
Qy      4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
Db    712 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 771

Qy     24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db    772 TTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAAAAGAAAGAACTTGAA------ 825

Qy     44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db    825 ------------------------------------------------------------ 825

Qy     64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db    826 ---------------AGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 870

Qy     84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db    871 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 930

Qy    104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db    931 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 984

Qy    124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db    985 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1044

Qy    144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db   1045 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104

Qy    164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db   1105 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1164

Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db   1165 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1224

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db   1225 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1284

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db   1285 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1344

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db   1345 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1404

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db   1405 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1464

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db   1465 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1524

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db   1525 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1584

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db   1585 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1644

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db   1645 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1704
```

```
Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db   1705 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 1764

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db   1765 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 1824

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   1825 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 1884

Qy    424 Gln-----------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
Db   1885 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 1944

Qy    438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 452
Db   1945 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2004

Qy    453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470
Db   2005 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2064

Qy    471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
Db   2065 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2124

Qy    491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db   2125 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2184

Qy    511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
Db   2185 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2244

Qy    530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db   2245 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2304

Qy    550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db   2305 ACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2364

Qy    570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
Db   2365 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2424

Qy    590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db   2425 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2484

Qy    610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db   2485 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2544

Qy    630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db   2545 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 2604

Qy    650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db   2605 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 2664

Qy    670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db   2665 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 2724

Qy    690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db   2725 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2784

Qy    710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db   2785 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 2844

Qy    730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db   2845 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 2904

Qy    750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db   2905 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 2964

Qy    770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db   2965 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3024

Qy    790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db   3025 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3084

Qy    810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db   3085 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3144

Qy    830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db   3145 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3204

Qy    850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db   3205 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3264

Qy    870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
Db   3265 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3324

Qy    890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db   3325 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3384

Qy    910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db   3385 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3444

Qy    930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db   3445 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3504

Qy    950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db   3505 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3564

Qy    970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db   3565 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 3624

Qy    990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db   3625 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 3684

Qy   1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db   3685 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 3744

Qy   1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
Db   3745 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 3804

Qy   1050 LysLeuValAsnAlaGlnSer 1056
Db   3805 TCAATGGTCAGTGTCCAGGCT 3825
```

RESULT 11
ABA94365

```
ID   ABA94365 standard; DNA; 3860 BP.
XX
AC   ABA94365;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human BCRP DNA related seq Id No. 1.
XX
KW   Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;
KW   hepatotropic; neurodegenerative; cytostatic; antianemic; muscular; BCRP;
KW   cardiant; gene therapy; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3843
FT                   /*tag=  a
XX
PN   WO200192877-A2.
XX
PD   06-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US17459.
XX
PR   31-MAY-2000; 2000US-0584586.
PR   29-MAY-2001; 2001US-0866866.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI   Sorrentino B,  Schuetz J;
XX
DR   WPI; 2002-114368/15.
DR   P-PSDB; ABB07266.
XX
PT   Identifying a stem cell, for treating e.g., muscular dystrophy,
PT   myocardial infarction, Parkinson's disease, or neurodegenerative
PT   disorders, comprises detecting the expression of an ATP transport
PT   protein (BCRP) by a cell  -
XX
PS   Disclosure; Page 53-55; 87pp; English.
XX
CC   The invention provides a method of identifying and/or isolating a stem
CC   cell that involves detecting the expression of an ATP transport protein
CC   containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
CC   comprising stem cells. The isolated stem cells may be used in the
CC   treatment of diseases such as muscular dystrophy, degenerative liver
CC   disorder, myocardial infarction, Parkinson's disease, degenerative
CC   disorders of the brain, and for tissue regeneration or replacement.
CC   Haematopoietic cells can be used in bone marrow transplants (e.g., for
CC   treatment of leukemia) and for ex vivo gene therapy for treating blood
CC   diseases such as sickle cell anemia and thalassemia. The stem cells can
CC   also be used as cell targets in gene therapy protocols. The present
CC   sequence represents a sequence related to the BCRP for which no relevant
CC   information has been provided in the specification.
XX
SQ   Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;

Alignment Scores:
Pred. No.:               1.88e-282     Length:          3860
Score:                   3100.50       Matches:         595
Percent Similarity:      75.54%        Conservative:    211
Best Local Similarity:   55.76%        Mismatches:      218
Query Match:             58.17%        Indels:          43
DB:                      24            Gaps:            6

US-09-873-409-4 (1-1058) x ABA94365 (1-3860)

Qy      4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
Db    712 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 771

Qy     24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db    772 TTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAAAAGAAAGAACTTGAA------ 825

Qy     44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db    825 ------------------------------------------------------------ 825

Qy     64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db    826 --------------AGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 870

Qy     84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db    871 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 930

Qy    104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db    931 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 984

Qy    124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db    985 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1044

Qy    144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db   1045 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104

Qy    164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db   1105 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1164

Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db   1165 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1224

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db   1225 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1284

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db   1285 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1344

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db   1345 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1404

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db   1405 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1464

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db   1465 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1524

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db   1525 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1584

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db   1585 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1644

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db   1645 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1704

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db   1705 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 1764

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db   1765 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 1824
```

```
Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   1825 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 1884

Qy    424 Gln------------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
Db   1885 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 1944

Qy    438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 452
Db   1945 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2004

Qy    453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470
Db   2005 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2064

Qy    471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
Db   2065 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2124

Qy    491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db   2125 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2184

Qy    511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
Db   2185 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2244

Qy    530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db   2245 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2304

Qy    550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db   2305 ACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2364

Qy    570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
Db   2365 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2424

Qy    590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db   2425 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2484

Qy    610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db   2485 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2544

Qy    630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db   2545 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 2604

Qy    650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db   2605 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 2664

Qy    670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db   2665 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 2724

Qy    690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db   2725 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2784

Qy    710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db   2785 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 2844

Qy    730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db   2845 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 2904

Qy    750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db   2905 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 2964

Qy    770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db   2965 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3024

Qy    790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db   3025 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3084

Qy    810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db   3085 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3144

Qy    830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db   3145 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3204

Qy    850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db   3205 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3264

Qy    870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
Db   3265 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3324

Qy    890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db   3325 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3384

Qy    910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db   3385 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3444

Qy    930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db   3445 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3504

Qy    950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db   3505 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3564

Qy    970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db   3565 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 3624

Qy    990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db   3625 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 3684

Qy   1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db   3685 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 3744

Qy   1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
Db   3745 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 3804

Qy   1050 LysLeuValAsnAlaGlnSer 1056
Db   3805 TCAATGGTCAGTGTCCAGGCT 3825
```

```
RESULT 12
ABA94366
ID   ABA94366 standard; DNA; 3860 BP.
XX
AC   ABA94366;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human BCRP DNA related seq Id No. 3.
```

```
XX
KW   Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;
KW   hepatotropic; neurodegenerative; cytostatic; antianemic; muscular; BCRP;
KW   cardiant; gene therapy; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3843
FT                   /*tag= a
XX
PN   WO200192877-A2.
XX
PD   06-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US17459.
XX
PR   31-MAY-2000; 2000US-0584586.
PR   29-MAY-2001; 2001US-0866866.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI   Sorrentino B,  Schuetz J;
XX
DR   WPI; 2002-114368/15.
DR   P-PSDB; ABB07267.
XX
PT   Identifying a stem cell, for treating e.g., muscular dystrophy,
PT   myocardial infarction, Parkinson's disease, or neurodegenerative
PT   disorders, comprises detecting the expression of an ATP transport
PT   protein (BCRP) by a cell  -
XX
PS   Disclosure; Page 59-60; 87pp; English.
XX
CC   The invention provides a method of identifying and/or isolating a stem
CC   cell that involves detecting the expression of an ATP transport protein
CC   containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
CC   comprising stem cells. The isolated stem cells may be used in the
CC   treatment of diseases such as muscular dystrophy, degenerative liver
CC   disorder, myocardial infarction, Parkinson's disease, degenerative
CC   disorders of the brain, and for tissue regeneration or replacement.
CC   Haematopoietic cells can be used in bone marrow transplants (e.g., for
CC   treatment of leukemia) and for ex vivo gene therapy for treating blood
CC   diseases such as sickle cell anemia and thalassemia. The stem cells can
CC   also be used as cell targets in gene therapy protocols. The present
CC   sequence represents a sequence related to the BCRP for which no relevant
CC   information has been provided in the specification.
XX
SQ   Sequence 3860 BP; 1135 A; 746 C; 956 G; 1023 T; 0 other;

Alignment Scores:
Pred. No.:              1.88e-282     Length:        3860
Score:                  3100.50       Matches:       595
Percent Similarity:     75.54%        Conservative:  211
Best Local Similarity:  55.76%        Mismatches:    218
Query Match:            58.17%        Indels:        43
DB:                     24            Gaps:          6

US-09-873-409-4 (1-1058) x ABA94366 (1-3860)

Qy      4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
Db    712 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 771

Qy     24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db    772 TTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAAAAGAAAGAACTTGAA------ 825

Qy     44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db    825 ------------------------------------------------------------ 825

Qy     64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db    826 ---------------AGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 870

Qy     84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db    871 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 930

Qy    104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db    931 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 984

Qy    124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db    985 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1044

Qy    144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db   1045 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104

Qy    164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db   1105 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1164

Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db   1165 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1224

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db   1225 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1284

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db   1285 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1344

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db   1345 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1404

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db   1405 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1464

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db   1465 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1524

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db   1525 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1584

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db   1585 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1644

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db   1645 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1704

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db   1705 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 1764

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db   1765 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 1824

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   1825 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 1884

Qy    424 Gln-----------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
```

```
Db   1885 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 1944

Qy    438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 452
Db   1945 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2004

Qy    453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470
Db   2005 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2064

Qy    471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
Db   2065 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2124

Qy    491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db   2125 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2184

Qy    511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
Db   2185 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2244

Qy    530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db   2245 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2304

Qy    550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db   2305 ACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2364

Qy    570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
Db   2365 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2424

Qy    590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db   2425 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2484

Qy    610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db   2485 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2544

Qy    630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db   2545 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 2604

Qy    650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db   2605 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 2664

Qy    670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db   2665 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 2724

Qy    690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db   2725 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2784

Qy    710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db   2785 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 2844

Qy    730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db   2845 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 2904

Qy    750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db   2905 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 2964

Qy    770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db   2965 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3024

Qy    790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db   3025 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3084

Qy    810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db   3085 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3144

Qy    830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db   3145 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3204

Qy    850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db   3205 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3264

Qy    870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
Db   3265 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3324

Qy    890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db   3325 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3384

Qy    910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db   3385 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3444

Qy    930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db   3445 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3504

Qy    950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db   3505 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3564

Qy    970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db   3565 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 3624

Qy    990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db   3625 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 3684

Qy   1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db   3685 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 3744

Qy   1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
Db   3745 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 3804

Qy   1050 LysLeuValAsnAlaGlnSer 1056
Db   3805 TCAATGGTCAGTGTCCAGGCT 3825
```

RESULT 13
AAF86127
ID   AAF86127 standard; cDNA; 4186 BP.
XX
AC   AAF86127;
XX
DT   25-JUN-2001 (first entry)
XX
DE   Cynomologous monkey P-glycoprotein cDNA.
XX
KW   Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   efflux pump; ss.
XX
OS   Macaca fascicularis.
XX
FH   Key   Location/Qualifiers

```
FT   CDS             100..3942
FT                   /*tag= a
FT                   /product= "PGP"
FT                   /note= "P-glycoprotein"
XX
PN   WO200123565-A1.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26592.
XX
PR   28-SEP-1999;   99US-0156921.
PR   12-OCT-1999;   99US-0158818.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-Crespi DT,  Crespi CL;
XX
DR   WPI; 2001-316136/33.
DR   P-PSDB; AAB81064.
XX
PT   Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein
PT   (PGP) and homologous PGP polypeptides are useful for predicting
PT   bioavailability of compound and increasing PGP transporter activity in
PT   cell  -
XX
PS   Example 1; Page 51-57; 84pp; English.
XX
CC   This invention relates to a polynucleotide sequence encoding a
CC   cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the
CC   PGP protein. PGP, also known as multidrug transporter, MDR1 is a member
CC   of the ABC transporter superfamily. The enzyme serves as an efflux pump
CC   exporting small molecules across the cell membrane. The invention
CC   includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence
CC   and protein, and also that of an allelic variant. The PGP polynucleotide
CC   sequence is useful for increasing PGP transporter activity in a cell.
CC   Antisense sequences of the cDNA are useful for inhibiting PGP transport
CC   activity in a mammalian cell. They may also be used for increasing the
CC   bioavailability of a drug. The present sequence represents cDNA encoding
CC   cynomologous monkey P-glycoprotein.
XX
SQ   Sequence 4186 BP; 1226 A; 799 C; 1039 G; 1122 T; 0 other;

Alignment Scores:
Pred. No.:              2.12e-282      Length:         4186
Score:                  3100.50        Matches:        594
Percent Similarity:     75.40%         Conservative:   212
Best Local Similarity:  55.57%         Mismatches:     216
Query Match:            58.17%         Indels:         47
DB:                     22             Gaps:           7

US-09-873-409-4 (1-1058) x AAF86127 (1-4186)

Qy      4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
Db    811 TCATTTACTGATAAAGAACTCTTAGCTTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 870

Qy     24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db    871 TTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAAAAGAAAGAACTCGAA------ 924

Qy     44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db    924 ------------------------------------------------------------ 924

Qy     64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db    925 ---------------AGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 969

Qy     84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db    970 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTTATCTATGCATCTTAT 1029

Qy    104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db   1030 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAAAGGAA------TATTCTATT 1083

Qy    124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db   1084 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1143

Qy    144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db   1144 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTTTGAAATCTTCAAGATA 1203

Qy    164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db   1204 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1263

Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db   1264 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1323

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db   1324 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1383

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db   1384 AGCGGCTGTGGGAAGAGCACAACGGTCCAGCTGATGCAGAGGCTTTATGACCCCACAGAG 1443

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db   1444 GGCATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAACGTAAGGTTTCTACGGGAA 1503

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db   1504 ATCATCGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1563

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db   1564 CGCTATGGTCGTGAAGATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1623

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db   1624 GCCTATGACTTTATCATGAAACTGCCTCAGAAATTTGACACCCTGGTTGGAGAGAGAGGG 1683

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db   1684 GCCCAGCTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1743

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db   1744 CCCAAGATCCTCCTGCTGGACGAGGCCACGTCAGCCTTGGACACAGAAAGTGAAGCAGTG 1803

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db   1804 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 1863

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db   1864 TTGTCTACGGTTCGTAATGCCGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 1923

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   1924 AAAGGAAATCATGATGAGCTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 1983

Qy    424 GlnAspIle--------------------LysLysAlaAspGlu---------GlnMet 433
Db   1984 CAGACAGCAGGAAATGAAATTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 2043

Qy    434 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu---------------Pro 448
Db   2044 GATACCTTGGAAATGTCTTCACATGATTCAGGATCCAGTCTAATAAGAAAAAGATCCACT 2103

Qy    449 LeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGln 468
```

```
Db   2104 CGTAGGAGTGTCCGTGGATCACAAGGCCAA------GACAGAAAGCTTAGTACCAAAGAG 2157

Qy    469 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 488
Db   2158 GCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACT 2217

Qy    489 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 508
Db   2218 GAGTGGCCTTATTTGTTGTTGGTGTATTTGTGCCATTATAAATGGAGGTCTGCAACCA 2277

Qy    509 ValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThr 527
Db   2278 GCATTTGCAGTAATATTTTCAAAGATTATAGGGATTTTTACAAGAAATGATGATGCCGAA 2337

Qy    528 ThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 547
Db   2338 ACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGTCCTTGGAATTGTTTCT 2397

Qy    548 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 567
Db   2398 TTTATTACATTTTTCCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAG 2457

Qy    568 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 587
Db   2458 CGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGCTGGTTTGATGAC 2517

Qy    588 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 607
Db   2518 CCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAA 2577

Qy    608 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 627
Db   2578 GGGGCTATAGGTTCCAGGCTTGCTATAATTACCCAGAATATAGCAAATCTTGGGACAGGA 2637

Qy    628 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 647
Db   2638 ATAATTATATCCTTAATCTATGGTTGGCAACTGACACTGTTACTCTTAGCAATTGTACCC 2697

Qy    648 ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp 667
Db   2698 ATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGAT 2757

Qy    668 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 687
Db   2758 AAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACT 2817

Qy    688 IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 707
Db   2818 GTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGATCAGAGTTTGCAGGTA 2877

Qy    708 GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis 727
Db   2878 CCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATCACGTTTTCCTTCACGCAG 2937

Qy    728 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla 747
Db   2938 GCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACAT 2997

Qy    748 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 767
Db   2998 AGTCTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATG 3057

Qy    768 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 787
Db   3058 GCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCAGCAGCC 3117

Qy    788 HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys 807
Db   3118 CACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACAGAAGGCCTA 3177

Qy    808 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys 827
Db   3178 AAGCCGAACACATTGGAAGGAAATGTCACATTTAATGAAGTTGTATTCAACTATCCCACC 3237

Qy    828 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 847
Db   3238 CGACTGGACATCCCAGTGCTTCAGGGGCTGAGCCTGGAAGTGAAGAAGGGCCAGACGCTG 3297

Qy    848 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 867
Db   3298 GCCCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACGGTGGTCCAGCTCCTGGAGCGGTTC 3357

Qy    868 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 887
Db   3358 TATGACCCCTTGGCGGGGAAAGTGCTGCTTGACGGCAAAGAAATAAAGCAACTGAATGTT 3417

Qy    888 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 907
Db   3418 CAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGC 3477

Qy    908 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 927
Db   3478 ATTAGTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTG 3537

Qy    928 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn 947
Db   3538 AGGGCAGCCAAGGAGGCCAATATACACGCCTTCATCGAGTCACTGCCTAATAAATATAGC 3597

Qy    948 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 967
Db   3598 ACCAGAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATA 3657

Qy    968 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 987
Db   3658 GCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACATCAGCTCTG 3717

Qy    988 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1007
Db   3718 GATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGTACC 3777

Qy   1008 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1027
Db   3778 TGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTT 3837

Qy   1028 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1047
Db   3838 CAGAATGGCAGAGTCAAGGAGCACGGCACACATCAGCAGCTGCTGGCACAGAAAGGCATC 3897

Qy   1048 TyrPheLysLeuValAsnAlaGlnSer 1056
Db   3898 TATTTTTCAATGGTCAGTGTCCAGGCT 3924
```

```
RESULT 14
AAF86128
ID   AAF86128 standard; cDNA; 4195 BP.
XX
AC   AAF86128;
XX
DT   25-JUN-2001  (first entry)
XX
DE   Cynomologous monkey P-glycoprotein cDNA variant 1.
XX
KW   Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   efflux pump; ss.
XX
OS   Macaca fascicularis.
XX
FH   Key             Location/Qualifiers
FT   CDS             100..3951
FT                   /*tag=  a
FT                   /product= "PGP"
FT                   /note= "P-glycoprotein"
FT   misc_feature    376..384
FT                   /*tag=  b
FT                   /note= "Insertion of 9 nucleotides relative to PGP
```

```
FT                     allelic variant AAF86127"
XX
PN   WO200123565-A1.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26592.
XX
PR   28-SEP-1999;   99US-0156921.
PR   12-OCT-1999;   99US-0158818.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-Crespi DT,  Crespi CL;
XX
DR   WPI; 2001-316136/33.
DR   P-PSDB; AAB81065.
XX
PT   Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein
PT   (PGP) and homologous PGP polypeptides are useful for predicting
PT   bioavailability of compound and increasing PGP transporter activity in
PT   cell  -
XX
PS   Example 1; Page 59-65; 84pp; English.
XX
CC   This invention relates to a polynucleotide sequence encoding a
CC   cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the
CC   PGP protein. PGP, also known as multidrug transporter, MDR1 is a member
CC   of the ABC transporter superfamily. The enzyme serves as an efflux pump
CC   exporting small molecules across the cell membrane. The invention
CC   includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence
CC   and protein, and also that of an allelic variant. The PGP polynucleotide
CC   sequence is useful for increasing PGP transporter activity in a cell.
CC   Antisense sequences of the cDNA are useful for inhibiting PGP transport
CC   activity in a mammalian cell. They may also be used for increasing the
CC   bioavailability of a drug. The present sequence represents cDNA encoding
CC   cynomologous monkey P-glycoprotein. This sequence contains a 9 nucleotide
CC   insert compared to the PGP allelic variant given in AAF86127.
XX
SQ   Sequence 4195 BP; 1230 A; 801 C; 1039 G; 1125 T; 0 other;

Alignment Scores:
Pred. No.:              2.13e-282     Length:         4195
Score:                  3100.50       Matches:        594
Percent Similarity:     75.40%        Conservative:   212
Best Local Similarity:  55.57%        Mismatches:     216
Query Match:            58.17%        Indels:         47
DB:                     22            Gaps:           7

US-09-873-409-4 (1-1058) x AAF86128 (1-4195)

Qy      4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
            |||   |||   |||||||||   |||||:::||||||||||||||||||||||||||
Db    820 TCATTTACTGATAAAGAACTCTTAGCTTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 879

Qy     24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
          |||::::::||||||||||||||||||||       |||:::||||||||:::
Db    880 TTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAAAAGAAAGAACTCGAA------ 933

Qy     44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63

Db    933 ------------------------------------------------------------ 933

Qy     64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
                           ||||||   :::||||||::::::||||||      ||||||||
Db    934 ---------------AGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 978

Qy     84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
          :::   |||   :::   :::|||:::||||||   :::       :::     :::|||
Db    979 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTTATCTATGCATCTTAT 1038

Qy    104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
             ||||||||||||||||||||:::|||:::|||:::   |||      |||:::|||
Db   1039 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAAAGGAA------TATTCTATT 1092

Qy    124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
          |||   ||||||   |||||||||||||||:::        ::::::   :::|||   |||
Db   1093 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1152

Qy    144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
             |||      |||   ||||||   ||||||||||||||||||   ||||||::::::
Db   1153 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTTTGAAATCTTCAAGATA 1212

Qy    164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
          ||||||   ||||||||||||||||:::::::|||   :::|||:::||||||::::::|||
Db   1213 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1272

Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
          :::|||   :::||||||:::||||||   |||:::||||||||||||      :::|||
Db   1273 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1332

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
          |||||||||||||||||||||||||:::::::::||||||:::||||||||||||||||||
Db   1333 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1392

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
          :::|||   ||||||||||||   |||||||||:::||||||||||||||||||   :::
Db   1393 AGCGGCTGTGGGAAGAGCACAACGGTCCAGCTGATGCAGAGGCTTTATGACCCCACAGAG 1452

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
          |||   :::   ||||||      ||||||||||   :::||||||||   |||:::
Db   1453 GGCATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAACGTAAGGTTTCTACGGGAA 1512

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
             ||||||||||||||||||||||||||||||   |||||||||:::   ||||||
Db   1513 ATCATCGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1572

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
          :::|||||||||:::||||||||   :::|||:::|||:::|||   :::|||||||||
Db   1573 CGCTATGGTCGTGAAGATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1632

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
          ||||||||||||||||:::   |||   ||||||:::|||||||||||||||:::|||
Db   1633 GCCTATGACTTTATCATGAAACTGCCTCAGAAATTTGACACCCTGGTTGGAGAGAGAGGG 1692

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
          ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1693 GCCCAGCTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1752

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
          ||||||||||:::||||||||||||||||||||||||||||||:::||||||::::::
Db   1753 CCCAAGATCCTCCTGCTGGACGAGGCCACGTCAGCCTTGGACACAGAAAGTGAAGCAGTG 1812

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
          ||||||   ||||||:::||||||   ||||||||||||||||||||:::|||||||||
Db   1813 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 1872

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
          |||||||||:::|||:::||||||:::|||               ||||||::::::   |||
Db   1873 TTGTCTACGGTTCGTAATGCCGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 1932

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
          ||||||   |||   |||||||||   ::::::|||:::|||:::   ||||||
Db   1933 AAAGGAAATCATGATGAGCTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 1992

Qy    424 GlnAspIle--------------------LysLysAlaAspGlu---------GlnMet 433
          |||                         |||||||||             ::::::
Db   1993 CAGACAGCAGGAAATGAAATTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 2052

Qy    434 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu---------------Pro 448
          :::::::::      |||:::         :::||||||
Db   2053 GATACCTTGGAAATGTCTTCACATGATTCAGGATCCAGTCTAATAAGAAAAAGATCCACT 2112

Qy    449 LeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGln 468
                |||||:::      :::         |||:::         :::   :::
Db   2113 CGTAGGAGTGTCCGTGGATCACAAGGCCAA------GACAGAAAGCTTAGTACCAAAGAG 2166
```

```
Qy     469 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 488
Db    2167 GCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACT 2226

Qy     489 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 508
Db    2227 GAGTGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGTCTGCAACCA 2286

Qy     509 ValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThr 527
Db    2287 GCATTTGCAGTAATATTTTCAAAGATTATAGGGATTTTTACAAGAAATGATGATGCCGAA 2346

Qy     528 ThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 547
Db    2347 ACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGTCCTTGGAATTGTTTCT 2406

Qy     548 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 567
Db    2407 TTTATTACATTTTTCCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAG 2466

Qy     568 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 587
Db    2467 CGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGCTGGTTTGATGAC 2526

Qy     588 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 607
Db    2527 CCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAA 2586

Qy     608 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 627
Db    2587 GGGGCTATAGGTTCCAGGCTTGCTATAATTACCCAGAATATAGCAAATCTTGGGACAGGA 2646

Qy     628 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 647
Db    2647 ATAATTATATCCTTAATCTATGGTTGGCAACTGACACTGTTACTCTTAGCAATTGTACCC 2706

Qy     648 ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp 667
Db    2707 ATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGAT 2766

Qy     668 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 687
Db    2767 AAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACT 2826

Qy     688 IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 707
Db    2827 GTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGATCAGAGTTTGCAGGTA 2886

Qy     708 GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis 727
Db    2887 CCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATCACGTTTTCCTTCACGCAG 2946

Qy     728 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla 747
Db    2947 GCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACAT 3006

Qy     748 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 767
Db    3007 AGTCTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATG 3066

Qy     768 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 787
Db    3067 GCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCAGCAGCC 3126

Qy     788 HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys 807
Db    3127 CACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACAGAAGGCCTA 3186

Qy     808 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys 827
Db    3187 AAGCCGAACACATTGGAAGGAAATGTCACATTTAATGAAGTTGTATTCAACTATCCCACC 3246

Qy     828 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 847
Db    3247 CGACTGGACATCCCAGTGCTTCAGGGGCTGAGCCTGGAAGTGAAGAAGGGCCAGACGCTG 3306

Qy     848 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 867
Db    3307 GCCCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACGGTGGTCCAGCTCCTGGAGCGGTTC 3366

Qy     868 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 887
Db    3367 TATGACCCCTTGGCGGGGAAAGTGCTGCTTGACGGCAAAGAAATAAAGCAACTGAATGTT 3426

Qy     888 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 907
Db    3427 CAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGC 3486

Qy     908 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 927
Db    3487 ATTAGTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTG 3546

Qy     928 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn 947
Db    3547 AGGGCAGCCAAGGAGGCCAATATACACGCCTTCATCGAGTCACTGCCTAATAAATATAGC 3606

Qy     948 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 967
Db    3607 ACCAGAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATA 3666

Qy     968 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 987
Db    3667 GCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACATCAGCTCTG 3726

Qy     988 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1007
Db    3727 GATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGTACC 3786

Qy    1008 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1027
Db    3787 TGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTT 3846

Qy    1028 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1047
Db    3847 CAGAATGGCAGAGTCAAGGAGCACGGCACACATCAGCAGCTGCTGGCACAGAAAGGCATC 3906

Qy    1048 TyrPheLysLeuValAsnAlaGlnSer 1056
Db    3907 TATTTTTCAATGGTCAGTGTCCAGGCT 3933
```

RESULT 15
AAV66533
ID   AAV66533 standard; cDNA; 4264 BP.
XX
AC   AAV66533;
XX
DT   08-JAN-1999 (first entry)
XX
DE   Mutated human P-glycoprotein gene designated mdr-delta-F335.
XX
KW   mdr1-delta-F335; cyclosporin-resistant deletion mutant; human;
KW   P-glycoprotein; cyclosporin; drug efflux;
KW   modulator-resistant multidrug resistance; cancer therapy; ss.
XX
OS   Homo sapiens.
XX
PN   US5830697-A.
XX
PD   03-NOV-1998.
XX
PF   21-JAN-1997;   97US-0784649.
XX
PR   21-JAN-1997;   97US-0784649.
XX
PA   (STRD) UNIV LELAND STANFORD JUNIOR.

```
XX
PI   Chen G,  Sikic BI;
XX
DR   WPI; 1998-609244/51.
XX
PT   DNA encoding mutant P-glycoprotein polypeptide - useful for
PT   producing recombinant polypeptide
XX
PS   Disclosure; Columns 15-20; 17pp; English.
XX
CC   The present sequence represents mdr1-delta-F335, a gene encoding a
CC   cyclosporin-resistant deletion mutant of human P-glycoprotein lacking
CC   Phe335. The recombinant protein can be used to produce antibodies. It
CC   can also be used in drug screening assays to study the effects of
CC   modulators such as cyclosporin on P-glycoprotein-mediated drug efflux.
CC   The nucleic acid can be used to transfer modulator-resistant multidrug
CC   resistance into transfected cells, e.g. haematopoietic stem cells for
CC   use during cancer therapy.
XX
SQ   Sequence 4264 BP; 1263 A; 807 C; 1041 G; 1153 T; 0 other;

Alignment Scores:
Pred. No.:              2.18e-282     Length:         4264
Score:                  3100.50       Matches:        595
Percent Similarity:     75.54%        Conservative:   211
Best Local Similarity:  55.76%        Mismatches:     218
Query Match:            58.17%        Indels:         43
DB:                     19            Gaps:           6

US-09-873-409-4 (1-1058) x AAV66533 (1-4264)

Qy       4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
             |||   |||   ||||||||  ||||||:::||||||||||||||||||||||||||
Db     850 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 909

Qy      24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
             |||::::::|||||||||||||||||||||     |||:::|||||||||:::
Db     910 TTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAAAAGAAAGAACTTGAA------ 963

Qy      44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63

Db     963 ------------------------------------------------------------ 963

Qy      64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
                           |||||   :::||||||::::::||||||      ||||||||
Db     964 ---------------AGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1008

Qy      84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
             :::    |||   :::   :::|||:::||||||   :::      :::      :::|||
Db    1009 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1068

Qy     104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
                |||||||||||||||||||:::|||:::|||:::||||||      |||:::|||
Db    1069 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1122

Qy     124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
             |||   ||||||   ||||||||||||||||:::       ::::::   :::|||   |||
Db    1123 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1182

Qy     144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
                |||      |||   ||||||      |||||||||||||:::     ||||||::::::
Db    1183 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1242

Qy     164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
             ||||||   ||||||||||||||||||::::::|||   :::|||:::|||||||::::::|||
Db    1243 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1302

Qy     184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
             :::|||   :::||||||:::||||||   |||:::||||||||||||      :::|||
Db    1303 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1362

Qy     204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
             ||||||||||||||||||||||:::::::::||||||:::||||||||||||||||||
Db    1363 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1422

Qy     224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
             :::|||   ||||||||||||   ||||||||:::||||||||||||||||||   :::
Db    1423 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1482

Qy     244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
             |||   :::   ||||||      ||||||||   :::|||||||||      ||||:::
Db    1483 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1542

Qy     264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
             |||||||||||||||||||||||||||||||||   |||||||||:::   ||||||
Db    1543 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1602

Qy     284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
             :::||||||||::::::||||||   :::|||:::|||:::|||   :::||||||||
Db    1603 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1662

Qy     304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
             |||||||||||||||||:::   |||:::||||||:::|||||||||||||||:::|||
Db    1663 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1722

Qy     324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
             ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1723 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1782

Qy     344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
             |||||||||||:::||||||||||||||||||||||||:::||||||::::::
Db    1783 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1842

Qy     364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
             ||||||   ||||||:::||||||   ||||||||||||||||||||:::||||||||
Db    1843 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 1902

Qy     384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
             |||||||||:::|||:::||||||:::|||             ||||||::::::   |||
Db    1903 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 1962

Qy     404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
             ||||||   |||   |||||||||   ::::::|||:::|||:::   ||||||
Db    1963 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 2022

Qy     424 Gln-----------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
             |||                 :::::::::   |||   ::::::   :::|||
Db    2023 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 2082

Qy     438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 452
                :::   |||   :::|||
Db    2083 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2142

Qy     453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470
                :::|||::::::   |||      |||:::      :::   ::::::
Db    2143 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2202

Qy     471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
             :::   |||:::|||   ||||||      :::|||:::||||||||      ||||||
Db    2203 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2262

Qy     491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
             |||:::   |||:::|||      :::::::::||||||   :::   |||   |||
Db    2263 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2322

Qy     511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
             :::||||||:::||||||||   :::|||      :::|||   |||
Db    2323 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2382

Qy     530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
             :::   ::::::   ::::::|||::::::|||:::   ||||||:::|||   |||:::
Db    2383 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2442

Qy     550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
             ::::::|||:::||||||      :::|||:::||||||||||||||||||   ||||||
Db    2443 ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2502
```

```
Qy      570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
            |||:::::::  |||:::::::||||||  ||||||::::::||||||||:::    :::
Db     2503 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2562

Qy      590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
            |||:::||||||  |||||||||  ||||||  |||  ||||||::::::||||||
Db     2563 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2622

Qy      610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
                ||||||||:::  |||:::|||||||||    |||:::|||     :::|||
Db     2623 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2682

Qy      630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
            |||||||||||||||||||||::::::|||  |||:::|||:::|||  |||::::::
Db     2683 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 2742

Qy      650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
            |||:::  |||::::::|||      ::::::|||  |||  |||||||||:::
Db     2743 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 2802

Qy      670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
            ||||||:::  ||||||||||||||||||||||||:::||||||  ||||||:::|||
Db     2803 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 2862

Qy      690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
            |||||||||:::|||:::  ||||||  ||||||  :::  ||||||      :::
Db     2863 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2922

Qy      710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
            ||||||:::  :::||||||  |||  |||    ::::::|||:::  |||
Db     2923 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 2982

Qy      730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
            :::||||||:::|||||||||  ||||||||||||||||||:::
Db     2983 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 3042

Qy      750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
            |||:::  |||   :::   :::||||||:::|||:::   :::||||||||||||:::
Db     3043 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 3102

Qy      770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
            |||:::            ||||||:::|||:::|||||||||   |||||||||:::
Db     3103 GGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3162

Qy      790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
               ::::::||||||  |||  |||||||||  |||  ||||||      |||
Db     3163 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3222

Qy      810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
            :::|||  |||||||||:::  |||  ||||||  |||  ||||||  ||||||
Db     3223 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3282

Qy      830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
            |||:::   :::|||:::||||||||||   :::::::::|||:::|||:::|||
Db     3283 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3342

Qy      850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
            |||||||||||||||||||||||||||   ||||||||||||:::|||   ||||||
Db     3343 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3402

Qy      870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
            |||:::  |||:::||||||  ||||||   :::  |||  |||||||||||||||
Db     3403 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3462

Qy      890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
            ||||||:::   :::   ||||||   |||||||||:::|||||||:::||||||||||||
Db     3463 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3522

Qy      910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
            |||||||||||||||||||||||||||||||||||      :::||||||      |||
Db     3523 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATTGTGAGGGCA 3582

Qy      930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
            |||      ||||||||||||:::|||||||||  ||||||  ||||||:::|||:::
Db     3583 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3642

Qy      950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
            ||||||  ||||||  |||||||||||||||||||||||||||:::||||||||||||
Db     3643 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3702

Qy      970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
            ||||||:::::::::|||   ||||||||||||||||||||||||||||||||||||
Db     3703 GCCCTTGTTAGACAGCCTCATATTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 3762

Qy      990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
            :::||||||||||||||||||   ||||||||||||||||||   ||||||||||||:::
Db     3763 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 3822

Qy     1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
            |||:::   ||||||||||||   ||||||||||||||||||||||||||||      |||
Db     3823 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 3882

Qy     1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
            |||::::::||||||   ||||||||||||:::||||||      :::   |||||||||
Db     3883 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 3942

Qy     1050 LysLeuValAsnAlaGlnSer 1056
               :::|||:::   |||:::
Db     3943 TCAATGGTCAGTGTCCAGGCT 3963
```

Search completed: March 30, 2003, 03:12:46
Job time : 1302.02 secs

```
                    GenCore version 5.1.4_p5_4578
           Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:        March 31, 2003, 02:24:20 ; Search time 154.939 Seconds
                                          (without alignments)
                                          2094.145 Million cell updates/sec

Title:         US-09-873-409-4
Perfect score: 5330
Sequence:      1 MVISLTSKELSAYSKAGAVA..........QELLRNRDIYFKLVNAQSVQ 1058

Scoring table: BLOSUM62
               Xgapop 10.0 , Xgapext  0.5
               Ygapop 10.0 , Ygapext  0.5
               Fgapop  6.0 , Fgapext  7.0
               Delop   6.0 , Delext   7.0

Searched:      441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters:      882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115422_19259/app_query.fasta_1.7544
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_578_@runat_27032003_115422_19259 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :     Issued_Patents_NA:*
               1:  /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
               2:  /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
               3:  /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
               4:  /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
               5:  /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
               6:  /cgn2_6/ptodata/2/ina/backfiles1.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3100.5	58.2	4264	2	US-08-784-649A-1	Sequence 1, Appli
2	3100.5	58.2	4264	2	US-08-784-649A-5	Sequence 5, Appli
3	3100.5	58.2	4646	1	US-08-181-471-2	Sequence 2, Appli
4	3100.5	58.2	4669	6	5206352-3	Patent No. 5206352
5	3100.5	58.2	6505	2	US-08-793-610-5	Sequence 5, Appli
6	3100.5	58.2	9318	2	US-08-793-610-6	Sequence 6, Appli
7	3086	57.9	4233	3	US-09-120-513-1	Sequence 1, Appli
8	3086	57.9	4233	4	US-09-450-105-1	Sequence 1, Appli
9	3073.5	57.7	4669	2	US-08-752-447-1	Sequence 1, Appli
10	3073.5	57.7	4669	4	US-09-316-167-1	Sequence 1, Appli
11	3056.5	57.3	4669	2	US-08-583-276-18	Sequence 18, Appl
12	2280	42.8	2726	1	US-08-461-823-1	Sequence 1, Appli
13	2067	38.8	4002	2	US-08-996-545-1	Sequence 1, Appli
14	2067	38.8	4002	2	US-08-996-545-3	Sequence 3, Appli
15	2067	38.8	4002	4	US-09-328-320-1	Sequence 1, Appli
16	2067	38.8	4002	4	US-09-328-320-3	Sequence 3, Appli
17	2038.5	38.2	4047	2	US-08-612-734B-1	Sequence 1, Appli
18	1995.5	37.4	4800	2	US-08-612-734B-3	Sequence 3, Appli
19	1926	36.1	3924	1	US-08-395-246C-1	Sequence 1, Appli
20	1885.5	35.4	4224	1	US-08-612-521-1	Sequence 1, Appli
21	1558	29.2	6143	1	US-08-612-521-3	Sequence 3, Appli
22	1546	29.0	3909	1	US-08-232-537-1	Sequence 1, Appli
23	1544.5	29.0	3924	2	US-08-996-644-3	Sequence 3, Appli
24	1544.5	29.0	3924	3	US-09-352-552-3	Sequence 3, Appli
25	1544.5	29.0	3927	2	US-08-996-644-1	Sequence 1, Appli
26	1544.5	29.0	3927	3	US-09-352-552-1	Sequence 1, Appli
27	1460.5	27.4	3792	4	US-09-351-224E-10	Sequence 10, Appl
28	1326	24.9	3999	4	US-09-351-224E-9	Sequence 9, Appli
29	1037	19.5	13188	4	US-08-961-527-70	Sequence 70, Appl
30	982.5	18.4	7760	4	US-08-961-527-63	Sequence 63, Appl
31	936.5	17.6	4403765	4	US-09-103-840A-2	Sequence 2, Appli
32	936.5	17.6	4411529	4	US-09-103-840A-1	Sequence 1, Appli
33	935.5	17.6	6492	4	US-08-961-527-188	Sequence 188, App
34	898	16.8	7186	4	US-08-961-527-39	Sequence 39, Appl
35	678	12.7	2376	1	US-08-394-880B-1	Sequence 1, Appli
36	659.5	12.4	5889	1	US-08-463-092B-5	Sequence 5, Appli
37	659.5	12.4	5889	2	US-08-462-109A-5	Sequence 5, Appli
38	659.5	12.4	5889	2	US-08-460-907B-5	Sequence 5, Appli
39	659.5	12.4	5889	3	US-08-463-179A-5	Sequence 5, Appli
40	659.5	12.4	5889	3	US-08-461-384B-5	Sequence 5, Appli
41	657	12.3	4781	2	US-09-001-273-1	Sequence 1, Appli
42	657	12.3	4781	4	US-08-843-459A-1	Sequence 1, Appli
43	657	12.3	4847	3	US-09-061-400-1	Sequence 1, Appli
44	649	12.2	5011	1	US-08-463-092B-3	Sequence 3, Appli
45	649	12.2	5011	2	US-08-462-109A-3	Sequence 3, Appli

ALIGNMENTS

```
RESULT 1
US-08-784-649A-1
; Sequence 1, Application US/08784649A
; Patent No. 5830697
;  GENERAL INFORMATION:
;    APPLICANT:  Sikic, Branimir I
;    APPLICANT:  Chen, Gang
;    TITLE OF INVENTION:  P-GLYCOPROTEIN MUTANT RESISTANT TO
;    TITLE OF INVENTION:  CYCLOSPORIN MODULATION
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  2200 Sand Hill Road
;      CITY:  Menlo Park
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/784,649A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sherwood, Pamela J
;      REGISTRATION NUMBER:  Reg.No. 5830697 36,677
;      REFERENCE/DOCKET NUMBER:  06037/007001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-322-5070
;      TELEFAX:  415-854-0875
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4264 base pairs
```

```
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-784-649A-1

Alignment Scores:
Pred. No.:                0              Length:          4264
Score:                    3100.50        Matches:         595
Percent Similarity:       75.54%         Conservative:    211
Best Local Similarity:    55.76%         Mismatches:      218
Query Match:              58.17%         Indels:          43
DB:                       2              Gaps:            6

US-09-873-409-4 (1-1058) x US-08-784-649A-1 (1-4264)

Qy       4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
             |||   |||   |||||||||   ||||||:::|||||||||||||||||||||||||
Db     850 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 909

Qy      24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
           |||::::::||||||||||||||||||||||||         |||:::|||||||||:::
Db     910 TTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAAAAGAAAGAACTTGAA------ 963

Qy      44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63

Db     963 ------------------------------------------------------------ 963

Qy      64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
                            ||||||   :::||||||::::::||||||      ||||||||
Db     964 ---------------AGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGATAAAG 1008

Qy      84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
           :::    |||    :::    :::|||:::||||||    :::         :::    :::|||
Db    1009 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1068

Qy     104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
              |||||||||||||||||||||:::||||||:::||||||         |||:::|||
Db    1069 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1122

Qy     124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
           |||   ||||||   ||||||||||||||||:::         ::::::    :::|||   |||
Db    1123 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1182

Qy     144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
              |||      |||   ||||||   |||||||||||||||:::   ||||||::::::
Db    1183 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1242

Qy     164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
           ||||||   ||||||||||||||||||::::::|||   :::|||:::|||||||::::::|||
Db    1243 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1302

Qy     184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
           :::|||   :::||||||:::||||||   |||:::|||||||||||||      :::|||
Db    1303 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1362

Qy     204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
           |||||||||||||||||||||||||:::::::::||||||:::||||||||||||||||||
Db    1363 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1422

Qy     224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
           :::|||   ||||||||||||   ||||||||:::||||||||||||||||||      :::
Db    1423 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1482

Qy     244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
           |||   :::   ||||||      |||||||||   :::||||||||||      |||:::
Db    1483 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1542

Qy     264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
              ||||||||||||||||||||||||||||||||      |||||||||:::      ||||||
Db    1543 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1602

Qy     284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
           :::||||||||::::::||||||   :::|||:::|||:::|||   :::|||||||||
Db    1603 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1662

Qy     304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
           ||||||||||||||||||:::   |||:::||||||:::||||||||||||||:::|||
Db    1663 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1722

Qy     324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
           ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1723 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1782

Qy     344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
           ||||||||||:::|||||||||||||||||||||||||||:::|||||||::::::
Db    1783 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1842

Qy     364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
           ||||||   ||||||:::||||||   |||||||||||||||||||||:::||||||||
Db    1843 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 1902

Qy     384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
           ||||||||:::|||:::||||||:::|||                  ||||||::::::      |||
Db    1903 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 1962

Qy     404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
           ||||||   |||   |||||||||.   ::::::|||::||||:::      ||||||
Db    1963 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 2022

Qy     424 Gln-----------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
           |||                  :::::::::      |||      ::::::      :::|||
Db    2023 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 2082

Qy     438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 452
              :::      |||         :::|||                  :::
Db    2083 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2142

Qy     453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470
              :::|||::::::      |||            |||:::               :::      ::::::
Db    2143 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2202

Qy     471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
           :::      |||:::|||   ||||||            :::|||:::||||||||            ||||||
Db    2203 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2262

Qy     491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
           |||:::      |||:::|||            :::::::::||||||      :::      |||   |||
Db    2263 CCTTATTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2322

Qy     511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
           :::||||||||:::||||||||      :::|||            :::|||      |||
Db    2323 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2382

Qy     530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
           :::   ::::::      ::::::|||::::::|||:::      ||||||:::|||      |||:::
Db    2383 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2442

Qy     550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
           ::::::|||:::||||||            :::|||:::||||||||||||||||||      ||||||
Db    2443 ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2502

Qy     570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
           |||::::::      |||::::::|||||||:::||||||||||||:::      :::
Db    2503 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2562

Qy     590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
           |||:::||||||   |||||||||      ||||||      |||      |||||||:::::::||||||
Db    2563 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2622

Qy     610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
              |||||||||:::      |||:::|||||||||            |||:::|||            :::|||
Db    2623 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2682

Qy     630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
           |||||||||||||||||||:::::::|||      |||:::|||:::|||      |||::::::
```

```
Db    2683 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 2742

Qy     650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
           |||:::   |||::::::|||        ::::::|||   |||   ||||||||:::
Db    2743 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 2802

Qy     670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
           ||||||:::   ||||||||||||||||||||||||:::||||||   ||||||:::|||
Db    2803 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 2862

Qy     690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
           |||||||||:::|||:::   ||||||   ||||||   :::   ||||||      :::
Db    2863 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2922

Qy     710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
           ||||||:::   :::||||||   |||   |||      ::::::|||:::   |||
Db    2923 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 2982

Qy     730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
           :::||||||:::||||||      ||||||||||||||||||||||||:::
Db    2983 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 3042

Qy     750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
           |||:::   |||. :::   :::||||||:::|||:::   :::||||||||||||:::
Db    3043 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 3102

Qy     770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
           |||:::            ||||||:::|||:::||||||||      |||||||||:::
Db    3103 GGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3162

Qy     790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
                 ::::::||||||   |||   |||||||||   |||   ||||||      |||
Db    3163 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3222

Qy     810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
           :::|||   |||||||||:::   |||   ||||||   |||   ||||||   ||||||
Db    3223 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3282

Qy     830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
           |||:::   :::|||:::||||||||||||   :::::::::|||:::|||:::|||
Db    3283 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3342

Qy     850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
           |||||||||||||||||||||||||||   ||||||||||||:::|||   ||||||
Db    3343 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3402

Qy     870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
           |||:::   |||:::||||||   ||||||   :::   |||   |||||||||||||||
Db    3403 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3462

Qy     890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
           ||||||:::   :::   ||||||   |||||||||:::||||||:::||||||||||||
Db    3463 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3522

Qy     910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
           ||||||||||||||||||||||||||||||||||||      :::||||||      |||
Db    3523 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATTGTGAGGGCA 3582

Qy     930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
           |||      |||||||||:::||||||||||   |||||||   ||||||:::|||:::
Db    3583 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3642

Qy     950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
           ||||||   ||||||   ||||||||||||||||||||||||||||:::||||||||||
Db    3643 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3702

Qy     970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
           ||||||:::::::::|||   ||||||||||||||||||||||||||||||||||||
Db    3703 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 3762

Qy     990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
           :::||||||||||||||||||   ||||||||||||||||||   ||||||||||||:::
Db    3763 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 3822

Qy    1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
           |||:::   ||||||||||||   ||||||||||||||||||||||||||||      |||
Db    3823 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 3882

Qy    1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
           |||::::::||||||   ||||||||||||:::||||||      :::   |||||||||
Db    3883 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 3942

Qy    1050 LysLeuValAsnAlaGlnSer 1056
              :::|||:::   |||:::
Db    3943 TCAATGGTCAGTGTCCAGGCT 3963

RESULT 2
US-08-784-649A-5
; Sequence 5, Application US/08784649A
; Patent No. 5830697
;  GENERAL INFORMATION:
;    APPLICANT:  Sikic, Branimir I
;    APPLICANT:  Chen, Gang
;    TITLE OF INVENTION:  P-GLYCOPROTEIN MUTANT RESISTANT TO
;    TITLE OF INVENTION:  CYCLOSPORIN MODULATION
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  2200 Sand Hill Road
;      CITY:  Menlo Park
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/784,649A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sherwood, Pamela J
;      REGISTRATION NUMBER:  Reg.No. 5830697 36,677
;      REFERENCE/DOCKET NUMBER:  06037/007001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-322-5070
;      TELEFAX:  415-854-0875
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4264 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-784-649A-5

Alignment Scores:
Pred. No.:                0              Length:        4264
Score:                    3100.50        Matches:       595
Percent Similarity:       75.54%         Conservative:  211
Best Local Similarity:    55.76%         Mismatches:    218
Query Match:              58.17%         Indels:        43
DB:                       2              Gaps:          6

US-09-873-409-4 (1-1058) x US-08-784-649A-5 (1-4264)

Qy       4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
           |||   |||   |||||||||      ||||||:::||||||||||||||||||||||||
Db     850 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 909

Qy      24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
           |||:::::::||||||||||||||||||         |||:::|||||||||:::
Db     910 TTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAAAAGAAAGAACTTGAA------ 963
```

```
Qy     44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db    963 ------------------------------------------------------------ 963
Qy     64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db    964 ---------------AGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1008
Qy     84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db   1009 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1068
Qy    104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db   1069 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1122
Qy    124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db   1123 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1182
Qy    144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db   1183 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1242
Qy    164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db   1243 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1302
Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db   1303 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1362
Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db   1363 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1422
Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db   1423 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1482
Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db   1483 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1542
Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db   1543 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1602
Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db   1603 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1662
Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db   1663 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1722
Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db   1723 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1782
Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db   1783 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1842
Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db   1843 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 1902
Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db   1903 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 1962
Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   1963 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 2022
Qy    424 Gln-----------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
Db   2023 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 2082
Qy    438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 452
Db   2083 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2142
Qy    453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470
Db   2143 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2202
Qy    471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
Db   2203 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2262
Qy    491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db   2263 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2322
Qy    511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
Db   2323 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2382
Qy    530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db   2383 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2442
Qy    550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db   2443 ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2502
Qy    570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
Db   2503 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2562
Qy    590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db   2563 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2622
Qy    610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db   2623 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2682
Qy    630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db   2683 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 2742
Qy    650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db   2743 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 2802
Qy    670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db   2803 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 2862
Qy    690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db   2863 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2922
Qy    710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db   2923 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 2982
Qy    730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db   2983 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 3042
Qy    750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
```

```
Db    3043 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 3102

Qy     770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
           |||:::              ||||||:::|||:::||||||||     |||||||||:::
Db    3103 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3162

Qy     790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
                 ::::::||||||     |||      |||||||||    |||    ||||||       |||
Db    3163 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3222

Qy     810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
           :::|||   |||||||||:::    |||    ||||||    |||    ||||||   ||||||
Db    3223 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3282

Qy     830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
           |||:::   :::|||:::||||||||||||     :::::::::|||:::|||:::|||
Db    3283 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3342

Qy     850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
           |||||||||||||||||||||||||||      |||||||||||:::|||   ||||||
Db    3343 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3402

Qy     870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
           |||:::   |||:::||||||   ||||||    :::    |||    ||||||||||||||
Db    3403 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3462

Qy     890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
           ||||||:::    :::    ||||||    |||||||||:::||||||:::||||||||||||
Db    3463 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3522

Qy     910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
           ||||||||||||||||||||||||||||||||||||      :::||||||      |||
Db    3523 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATTGTGAGGGCA 3582

Qy     930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
           |||      |||||||||||:::|||||||||    ||||||    ||||||:::|||:::
Db    3583 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3642

Qy     950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
           ||||||   ||||||   |||||||||||||||||||||||||||:::||||||||||||
Db    3643 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3702

Qy     970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
           ||||||::::::::|||   ||||||||||||||||||||||||||||||||||||
Db    3703 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 3762

Qy     990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
           :::||||||||||||||||||   |||||||||||||||||||   ||||||||||||:::
Db    3763 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 3822

Qy    1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
           |||:::   ||||||||||||   ||||||||||||||||||||||||||||||      |||
Db    3823 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 3882

Qy    1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
           |||::::::||||||   ||||||||||||:::||||||      :::   |||||||||
Db    3883 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 3942

Qy    1050 LysLeuValAsnAlaGlnSer 1056
                 :::|||:::   |||:::
Db    3943 TCAATGGTCAGTGTCCAGGCT 3963
```

RESULT 3
US-08-181-471-2
; Sequence 2, Application US/08181471
; Patent No. 5641508
; GENERAL INFORMATION:
;   APPLICANT: Li, Lingna
;   APPLICANT: Lishko, Valeryi K.
;   TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL
;   TITLE OF INVENTION: COMPOSITIONS TO HAIR FOLLICLES
;   NUMBER OF SEQUENCES: 3
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Thomas Fitting
;   STREET: 12526 High Bluff Drive, Suite 300
;   CITY: San Diego
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 92130
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/181,471
;   FILING DATE: 13-JAN-1994
;   CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/041,553
;   FILING DATE: 02-APR-1993
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Fitting, Thomas
;   REGISTRATION NUMBER: 34,163
;   REFERENCE/DOCKET NUMBER: ANT0029P
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 619-792-3680
;   TELEFAX: 619-792-8477
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4646 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 425..4267
US-08-181-471-2

```
Alignment Scores:
Pred. No.:              0              Length:         4646
Score:                  3100.50        Matches:        595
Percent Similarity:     75.54%         Conservative:   211
Best Local Similarity:  55.76%         Mismatches:     218
Query Match:            58.17%         Indels:         43
DB:                     1              Gaps:           6

US-09-873-409-4 (1-1058) x US-08-181-471-2 (1-4646)

Qy       4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
              |||    |||    |||||||||    ||||||:::||||||||||||||||||||||||
Db    1136 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 1195

Qy      24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
           |||:::::::||||||||||||||||||||||      |||:::|||||||||:::
Db    1196 TTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAAAAGAAAGAACTTGAA------ 1249

Qy      44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63

Db    1249 ------------------------------------------------------------ 1249

Qy      64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
                            ||||||   :::||||||:::::::||||||      |||||||||
Db    1250 ---------------AGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1294

Qy      84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
           :::   |||    :::    :::|||:::||||||         :::         :::|||
Db    1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1354

Qy     104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
              ||||||||||||||||||:::|||:::|||:::||||||       |||:::|||
Db    1355 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1408
```

```
Qy    124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db   1409 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1468

Qy    144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db   1469 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528

Qy    164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db   1529 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1588

Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db   1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1648

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db   1649 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1708

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db   1709 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1768

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db   1769 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1828

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db   1829 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1888

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db   1889 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1948

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db   1949 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2008

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db   2009 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2068

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db   2069 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2128

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db   2129 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 2188

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db   2189 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 2248

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   2249 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 2308

Qy    424 Gln------------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
Db   2309 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 2368

Qy    438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 452
Db   2369 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2428

Qy    453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470
Db   2429 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2488

Qy    471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
Db   2489 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2548

Qy    491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db   2549 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2608

Qy    511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
Db   2609 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2668

Qy    530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db   2669 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2728

Qy    550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db   2729 ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2788

Qy    570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
Db   2789 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2848

Qy    590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db   2849 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2908

Qy    610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db   2909 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2968

Qy    630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db   2969 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 3028

Qy    650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db   3029 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 3088

Qy    670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db   3089 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 3148

Qy    690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db   3149 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 3208

Qy    710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db   3209 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 3268

Qy    730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db   3269 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 3328

Qy    750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db   3329 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 3388

Qy    770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db   3389 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3448

Qy    790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db   3449 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3508

Qy    810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db   3509 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3568

Qy    830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
```

```
Db    3569 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3628

Qy     850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
           |||||||||||||||||||||||||||||    ||||||||||:::|||  |||||
Db    3629 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3688

Qy     870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
           |||:::   |||:::||||||  ||||||  :::  |||  ||||||||||||||
Db    3689 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3748

Qy     890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
           ||||||:::  :::  ||||||  |||||||||:::||||||:::||||||||||||
Db    3749 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3808

Qy     910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
           |||||||||||||||||||||||||||||||||||      :::||||||    |||
Db    3809 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3868

Qy     930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
           |||      ||||||||||||:::|||||||||  ||||||  ||||||:::|||:::
Db    3869 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3928

Qy     950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
           ||||||  ||||||  ||||||||||||||||||||||||||||:::||||||||||||
Db    3929 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3988

Qy     970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
           ||||||:::::::::|||  |||||||||||||||||||||||||||||||||
Db    3989 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 4048

Qy     990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
           :::|||||||||||||||||||  ||||||||||||||||||  ||||||||||||:::
Db    4049 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 4108

Qy    1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
           ||||:::  ||||||||||||  ||||||||||||||||||||||||||||   |||
Db    4109 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 4168

Qy    1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
           ||||::::::||||||  ||||||||||||:::||||||    :::  |||||||||
Db    4169 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 4228

Qy    1050 LysLeuValAsnAlaGlnSer 1056
              :::|||:::  |||:::
Db    4229 TCAATGGTCAGTGTCCAGGCT 4249
```

RESULT 4
5206352-3
;Patent No. 5206352
; APPLICANT: Roninson, Igor B.;Pastan Ira H.;Gottesman,
;Michael M.
; TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA
;SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/622,836
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 892,575
; FILING DATE: 01-AUG-1986
; APPLICATION NUMBER: 845,610
; FILING DATE: 28-MAR-1986
;SEQ ID NO:3:
; LENGTH: 4669
5206352-3

Alignment Scores:

Pred. No.:	0	Length:	4669
Score:	3100.50	Matches:	595
Percent Similarity:	75.54%	Conservative:	211
Best Local Similarity:	55.76%	Mismatches:	218
Query Match:	58.17%	Indels:	43
DB:	6	Gaps:	6

US-09-873-409-4 (1-1058) x 5206352-3 (1-4669)

```
Qy       4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
           |||  |||  |||||||||  |||||:::|||||||||||||||||||||||||||||
Db    1136 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 1195

Qy      24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
           |||::::::|||||||||||||||||||||      |||:::|||||||||:::
Db    1196 TTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAAAAGAAAGAACTTGAA------ 1249

Qy      44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63

Db    1249 ------------------------------------------------------------ 1249

Qy      64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
                          ||||||   :::|||||||::::::||||||    |||||||||
Db    1250 ---------------AGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1294

Qy      84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
           :::    |||    :::    :::|||:::||||||    :::      :::   :::|||
Db    1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1354

Qy     104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
           ||||||||||||||||||||||||||||||:::|||:::|||||||||    |||:::|||
Db    1355 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1408

Qy     124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
           |||   ||||||   ||||||||||||||||:::      ::::::   :::|||   |||
Db    1409 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1468

Qy     144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
              |||      |||   ||||||   ||||||||||||||:::   ||||||::::::
Db    1469 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528

Qy     164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
           ||||||   |||||||||||||||:::::::|||   :::|||:::||||||::::::|||
Db    1529 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1588

Qy     184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
           :::|||   :::|||||:::|||||||   |||:::||||||||||||      :::|||
Db    1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1648

Qy     204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
           |||||||||||||||||||||||:::::::::||||||:::||||||||||||||||||
Db    1649 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1708

Qy     224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
           :::|||   ||||||||||||   ||||||||:::|||||||||||||||||||   :::
Db    1709 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1768

Qy     244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
           |||   :::   ||||||      |||||||||   :::|||||||||      |||:::
Db    1769 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1828

Qy     264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
              ||||||||||||||||||||||||||||||||||||||||   |||:::   ||||||
Db    1829 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1888

Qy     284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
           :::||||||||||::::::||||||   :::|||:::|||:::|||   :::|||||||||
Db    1889 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1948

Qy     304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
           |||||||||||||||||:::   |||:::||||||:::|||||||||||||||:::|||
Db    1949 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2008

Qy     324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
           ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2009 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2068

Qy     344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
           ||||||||||||:::||||||||||||||||||||||||||:::||||||||::::::
```

```
Db   2069 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2128

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db   2129 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 2188

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db   2189 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 2248

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   2249 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 2308

Qy    424 Gln------------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
Db   2309 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 2368

Qy    438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 452
Db   2369 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2428

Qy    453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470
Db   2429 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2488

Qy    471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
Db   2489 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2548

Qy    491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db   2549 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2608

Qy    511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
Db   2609 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2668

Qy    530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db   2669 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2728

Qy    550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db   2729 ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2788

Qy    570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
Db   2789 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2848

Qy    590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db   2849 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2908

Qy    610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db   2909 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2968

Qy    630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db   2969 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 3028

Qy    650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db   3029 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 3088

Qy    670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db   3089 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 3148

Qy    690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db   3149 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 3208

Qy    710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db   3209 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 3268

Qy    730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db   3269 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 3328

Qy    750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db   3329 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 3388

Qy    770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db   3389 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3448

Qy    790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db   3449 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3508

Qy    810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db   3509 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3568

Qy    830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db   3569 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3628

Qy    850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db   3629 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3688

Qy    870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
Db   3689 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3748

Qy    890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db   3749 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3808

Qy    910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db   3809 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3868

Qy    930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db   3869 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3928

Qy    950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db   3929 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3988

Qy    970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db   3989 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 4048

Qy    990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db   4049 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 4108

Qy   1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db   4109 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 4168

Qy   1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
Db   4169 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 4228

Qy   1050 LysLeuValAsnAlaGlnSer 1056
Db   4229 TCAATGGTCAGTGTCCAGGCT 4249
```

```
RESULT 5
US-08-793-610-5
; Sequence 5, Application US/08793610
; Patent No. 5858744
;  GENERAL INFORMATION:
;    APPLICANT:  BAUM, Christopher
;    APPLICANT:  STOCKING-HARBERS, Carol
;    APPLICANT:  OSTERTAG, Wolfram
;    TITLE OF INVENTION:  RETROVIRAL VECTOR HYBRIDS AND THE USE THEROF
;    TITLE OF INVENTION:  FOR GENE TRANSFER
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nikaido, Marmelstein, Murray & Oram LLP
;      STREET:  655 Fifteenth Street N.W. Suite 330
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/793,610
;      FILING DATE:  07-MAR-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P 44 31 973.8
;      FILING DATE:  08-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 195 03 952.1
;      FILING DATE:  07-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/03175
;      FILING DATE:  10-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berman, Richard J.
;      REGISTRATION NUMBER:  39,105
;      REFERENCE/DOCKET NUMBER:  P1614-7007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)638-5000
;      TELEFAX:  (202)638-4810
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6505 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA
US-08-793-610-5

Alignment Scores:
Pred. No.:              0            Length:         6505
Score:                  3100.50      Matches:        595
Percent Similarity:     75.54%       Conservative:   211
Best Local Similarity:  55.76%       Mismatches:     218
Query Match:            58.17%       Indels:         43
DB:                     2            Gaps:           6

US-09-873-409-4 (1-1058) x US-08-793-610-5 (1-6505)

Qy      4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
Db   2528 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 2587

Qy     24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db   2588 TTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAAAAGAAAGAACTTGAA------ 2641

Qy     44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db   2641 ------------------------------------------------------------ 2641

Qy     64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db   2642 ---------------AGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 2686

Qy     84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db   2687 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 2746

Qy    104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db   2747 GCTCTGGCCTTCTGGTATGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 2800

Qy    124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db   2801 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 2860

Qy    144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db   2861 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 2920

Qy    164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db   2921 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 2980

Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db   2981 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 3040

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db   3041 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 3100

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db   3101 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 3160

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db   3161 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 3220

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db   3221 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 3280

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db   3281 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 3340

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db   3341 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 3400

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db   3401 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 3460

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db   3461 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 3520

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db   3521 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 3580

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db   3581 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 3640

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   3641 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 3700

Qy    424 Gln-----------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
```

```
Db   3701 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 3760

Qy    438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 452
Db   3761 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 3820

Qy    453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470
Db   3821 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 3880

Qy    471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
Db   3881 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 3940

Qy    491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db   3941 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 4000

Qy    511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
Db   4001 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 4060

Qy    530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db   4061 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 4120

Qy    550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db   4121 ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 4180

Qy    570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
Db   4181 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 4240

Qy    590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db   4241 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 4300

Qy    610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db   4301 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 4360

Qy    630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db   4361 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 4420

Qy    650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db   4421 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 4480

Qy    670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db   4481 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 4540

Qy    690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db   4541 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 4600

Qy    710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db   4601 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 4660

Qy    730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db   4661 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 4720

Qy    750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db   4721 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 4780

Qy    770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db   4781 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 4840

Qy    790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db   4841 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 4900

Qy    810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db   4901 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 4960

Qy    830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db   4961 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 5020

Qy    850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db   5021 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 5080

Qy    870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
Db   5081 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 5140

Qy    890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db   5141 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 5200

Qy    910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db   5201 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 5260

Qy    930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db   5261 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 5320

Qy    950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db   5321 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 5380

Qy    970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db   5381 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 5440

Qy    990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db   5441 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 5500

Qy   1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db   5501 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 5560

Qy   1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
Db   5561 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 5620

Qy   1050 LysLeuValAsnAlaGlnSer 1056
Db   5621 TCAATGGTCAGTGTCCAGGCT 5641
```

RESULT 6
US-08-793-610-6
; Sequence 6, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
; APPLICANT: BAUM, Christopher
; APPLICANT: STOCKING-HARBERS, Carol
; APPLICANT: OSTERTAG, Wolfram
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEROF
; TITLE OF INVENTION: FOR GENE TRANSFER
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington

```
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/793,610
;      FILING DATE:  07-MAR-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P 44 31 973.8
;      FILING DATE:  08-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 195 03 952.1
;      FILING DATE:  07-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/03175
;      FILING DATE:  10-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berman, Richard J.
;      REGISTRATION NUMBER:  39,105
;      REFERENCE/DOCKET NUMBER:  P1614-7007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)638-5000
;      TELEFAX:  (202)638-4810
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9318 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA
US-08-793-610-6

Alignment Scores:
Pred. No.:              0           Length:          9318
Score:                  3100.50     Matches:         595
Percent Similarity:     75.54%      Conservative:    211
Best Local Similarity:  55.76%      Mismatches:      218
Query Match:            58.17%      Indels:          43
DB:                     2           Gaps:            6

US-09-873-409-4 (1-1058) x US-08-793-610-6 (1-9318)

Qy      4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
Db   2487 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 2546

Qy     24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db   2547 TTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAAAAGAAAGAACTTGAA------ 2600

Qy     44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db   2600 ------------------------------------------------------------ 2600

Qy     64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db   2601 ---------------AGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 2645

Qy     84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db   2646 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 2705

Qy    104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db   2706 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 2759

Qy    124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db   2760 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 2819

Qy    144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db   2820 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 2879

Qy    164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db   2880 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 2939

Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db   2940 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 2999

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db   3000 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 3059

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db   3060 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 3119

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db   3120 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 3179

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db   3180 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 3239

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db   3240 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 3299

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db   3300 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 3359

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db   3360 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 3419

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db   3420 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 3479

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db   3480 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 3539

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db   3540 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 3599

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   3600 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 3659

Qy    424 Gln-----------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
Db   3660 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 3719

Qy    438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 452
Db   3720 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 3779

Qy    453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470
Db   3780 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 3839

Qy    471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
Db   3840 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 3899
```

```
Qy    491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db   3900 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 3959

Qy    511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
Db   3960 GCAATAATATTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 4019

Qy    530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db   4020 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 4079

Qy    550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db   4080 ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 4139

Qy    570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
Db   4140 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 4199

Qy    590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db   4200 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 4259

Qy    610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db   4260 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 4319

Qy    630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db   4320 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 4379

Qy    650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db   4380 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 4439

Qy    670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db   4440 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 4499

Qy    690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db   4500 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 4559

Qy    710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db   4560 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 4619

Qy    730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db   4620 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 4679

Qy    750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db   4680 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 4739

Qy    770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db   4740 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 4799

Qy    790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db   4800 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 4859

Qy    810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db   4860 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 4919

Qy    830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db   4920 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 4979

Qy    850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db   4980 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 5039

Qy    870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
Db   5040 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 5099

Qy    890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db   5100 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 5159

Qy    910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db   5160 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 5219

Qy    930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db   5220 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 5279

Qy    950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db   5280 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 5339

Qy    970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db   5340 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 5399

Qy    990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db   5400 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 5459

Qy   1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db   5460 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 5519

Qy   1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
Db   5520 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 5579

Qy   1050 LysLeuValAsnAlaGlnSer 1056
Db   5580 TCAATGGTCAGTGTCCAGGCT 5600
```

RESULT 7
US-09-120-513-1
; Sequence 1, Application US/09120513
; Patent No. 6025160
; GENERAL INFORMATION:
; APPLICANT: Brun, Kimberly
; APPLICANT: Chenery, Richard
; APPLICANT: Ellens, Harma
; APPLICANT: Field, John
; APPLICANT: Yue, Lin
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE
; TITLE OF INVENTION: SEQUENCES ENCODING RAT MDR1B2 AND
; TITLE OF INVENTION: SCREENING METHODS THEREOF
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY:
; ZIP: 19406
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/120,513
; FILING DATE: 22-JUL-1998
; CLASSIFICATION:

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  King, William T
;      REGISTRATION NUMBER:  30,954
;      REFERENCE/DOCKET NUMBER:  GP50008
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-5015
;      TELEFAX:  610-270-5090
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4233 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
US-09-120-513-1

Alignment Scores:
Pred. No.:               0          Length:        4233
Score:                   3086.00    Matches:       601
Percent Similarity:      74.74%     Conservative:  198
Best Local Similarity:   56.22%     Mismatches:    224
Query Match:             57.90%     Indels:        46
DB:                      3          Gaps:          7

US-09-873-409-4 (1-1058) x US-09-120-513-1 (1-4233)

Qy       1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20
Db     722 GTACTGACTTCATTTACTAATAAGGAACTCCAGGCTTATGCGAAAGCTGGAGCAGTTGCC 781

Qy      21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40
Db     782 GAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCGTTTGGAGGACAAAAGAAGGAACTT 841

Qy      41 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60
Db     842 GAA--------------------------------------------------------- 844

Qy      61 LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80
Db     845 ------------------------AGGTACAATAAAAATTTAGAAGAAGCTAAAAGAGTT 880

Qy      81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 100
Db     881 GGCATAAAGAAAGCCATCACGGCCAACATTTCCATAGGTATTGCCTACCTGTTGGTCTAT 940

Qy     101 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120
Db     941 GCGTCTTATGCACTGGCATTCTGGTATGGGACCTCCTTGGTCCTCTCAAATGAA------ 994

Qy     121 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 140
Db     995 TATTCTATTGGACAAGTGCTTACCGTCTTCTTCTCTATTTTATTGGGGACTTTCAGTATT 1054

Qy     141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
Db    1055 GGACATTTAGCCCCAAACATAGAAGCCTTTGCAAATGCAAGAGGGGCAGCCTATGAAATC 1114

Qy     161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180
Db    1115 TTCAAGATAATTGATAATGAGCCAAGCATTGACAGCTTCTCAACCAAGGGACACAAACCA 1174

Qy     181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200
Db    1175 GACAGTATAATGGGAAATTTGGAATTTAAAAATGTTTACTTCAACTACCCATCACGAAGT 1234

Qy     201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
Db    1235 GAAGTTAAGATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTG 1294

Qy     221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240
Db    1295 GTTGGCAACAGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGAC 1354

Qy     241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260
Db    1355 CCCATAGAGGGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATCAATGTGAGGTAT 1414

Qy     261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280
Db    1415 CTGCGGGAAATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCCACCACGATTGCC 1474

Qy     281 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 300
Db    1475 GAAAACATTCGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAG 1534

Qy     301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320
Db    1535 GAAGCCAATGCCTATGACTTCATCATGAAACTGCCCCACAAATTTAACACCCTGGTTGGT 1594

Qy     321 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340
Db    1595 GAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTG 1654

Qy     341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360
Db    1655 GTCCGCAACCCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGC 1714

Qy     361 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380
Db    1715 GAAGCCGTGGTTCAGGCCGCTCTGGATAAGGCTAGAGAAGGCCGGACCACCATTGTGATA 1774

Qy     381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400
Db    1775 GCTCACCGCTTGTCTACAGTGCGCAATGCTGACGTCATTGCTGGTTTTGATGGTGGTGTC 1834

Qy     401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420
Db    1835 ATTGTGGAGCAAGGAAATCATGAAGAGCTCATGAAAGAGAAGGGCATTTACTTCAAACTT 1894

Qy     421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440
Db    1895 GTCATGACACAG---ACTAGAGGAAATGAAATTGAACCAGGAAATAATGCTTATGAATCC 1951

Qy     441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460
Db    1952 CAAAGTGACACTGGTGCCTCTGAGTTGACTTCAGAAAAATCA---AAATCTCCTTTAATA 2008

Qy     461 AspLysAla--------------------------------GluGluSerThrGlnSer 469
Db    2009 AGGAGATCAATTCGCAGAAGTATCCACAGAAGACAAGACCAGGAGAGAAGACTTAGTTCG 2068

Qy     470 LysGlu---------IleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 486
Db    2069 AAAGAGGATGTGGATGAAGATGTGCCTATGGTTTCCTTTTGGCAGATCCTAAAGCTAAAT 2128

Qy     487 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 506
Db    2129 ATTAGTGAATGGCCCTATTTAGTTGTGGGTGTACTTTGTGCTGTTATAAATGGGTGCATA 2188

Qy     507 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGly---AsnAsnAsp 525
Db    2189 CAACCAGTGTTTGCCATAGTGTTTTCAAAGATTGTAGGGGTTTTTTCAAGAGACGACGAC 2248

Qy     526 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 545
Db    2249 CATGAAACCAAACAACGGAATTGTAACTTGTTTTCCCTTCTCTTTCTGGTCATGGGAATG 2308

Qy     546 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 565
Db    2309 ATTTCTTTTGTTACGTACTTCTTTCAAGGCTTCACATTTGGCAAAGCTGGAGAGATCCTC 2368

Qy     566 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 585
Db    2369 ACCAAGCGACTCCGATACATGGTCTTCAAATCCATGCTGCGACAGGATATAAGCTGGTTT 2428
```

```
Qy    586 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 605
Db   2429 GATGACCATAAAAACACCACTGGCTCGCTGACTACCAGGCTCGCTAGTGACGCTTCTAAT 2488

Qy    606 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 625
Db   2489 GTTAAAGGGGCTATGGGCTCCAGGCTTGCTGTAGTTACCCAGAATGTAGCAAACCTTGGC 2548

Qy    626 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 645
Db   2549 ACAGGAATTATCTTATCCTTAGTCTATGGCTGGCAGCTTACACTTTTACTTGTAGTAATT 2608

Qy    646 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 665
Db   2609 ATACCACTCATTGTCTTGGGTGGAATTATTGAAATGAAACTGTTGTCTGGTCAAGCCTTG 2668

Qy    666 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 685
Db   2669 AAGGACAAGAAAGAGCTAGAGATCTCTGGGAAGATCGCTACAGAAGCAATTGAAAACTTC 2728

Qy    686 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 705
Db   2729 CGCACTGTTGTCTCTTTGACTAGGGAGCAGAAGTTTGAAACTATGTATGCCCAGAGCTTG 2788

Qy    706 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 725
Db   2789 CAGATACCATACAGAAATGCTTTGAAGAAAGCACACGTCTTTGGGATCACCTTCGCCTTC 2848

Qy    726 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 745
Db   2849 ACCCAGGCCATGATTTATTTTTCCTATGCTGCTTGTTTCCGGTTCGGTGCCTACTTGGTG 2908

Qy    746 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 765
Db   2909 GCACGAGAACTCATGACGTTTGAAAATGTTATGTTGGTATTTTCTGCTGTTGTCTTTGGT 2968

Qy    766 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 785
Db   2969 GCCATGGCAGCAGGGAATACCAGTTCATTCGCTCCTGACTACGCGAAGGCCAAAGTCTCA 3028

Qy    786 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 805
Db   3029 GCATCCCACATCATCAGGATCATTGAGAAAATCCCCGAGATTGACAGCTACAGCACGGAG 3088

Qy    806 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 825
Db   3089 GGCTTGAAGCCTAATTGGTTAGAAGGAAATGTGAAATTTAATGGAGTCATGTTCAACTAT 3148

Qy    826 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 845
Db   3149 CCCACCCGACCCAACATCCCAGTGCTTCAGGGACTGAGCTTCGAGGTGAAGAAGGGGCAA 3208

Qy    846 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 865
Db   3209 ACGCTTCGCCTGGTGGGCAGCAGTGGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTCGAG 3268

Qy    866 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 885
Db   3269 CGCTTCTACAACCCCATGGCTGGAACAGTGTTTCTAGATGGCAAAGAAATAAAACAACTC 3328

Qy    886 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 905
Db   3329 AACGTCCAGTGGCTCCGCGCCCACCTGGGCATTGTGTCCCAGGAGCCCATCCTGTTTGAC 3388

Qy    906 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 925
Db   3389 TGCAGCATCACCGAGAACATCGCCTACGGAGACAACAGCCGTGTCGTGTCTCATGAGGAG 3448

Qy    926 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 945
Db   3449 ATCGTGAGGGCCGCCAGGGAGGCCAACATCCACCAGTTCATCGACTCACTGCCTGAGAAA 3508

Qy    946 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 965
Db   3509 TACAACACCAGAGTGGGAGACAAAGGGACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATC 3568

Qy    966 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 985
Db   3569 GCCATCGCGCGCGCCCTCGTCAGACAGCCTCACATCTTACTTCTGGATGAAGCGACATCA 3628

Qy    986 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1005
Db   3629 GCTCTGGATACGGAGAGTGAAAAGGTCGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGC 3688

Qy   1006 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1025
Db   3689 CGCACCTGCATTGTGATCGCGCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTG 3748

Qy   1026 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1045
Db   3749 GTGATTCAGAACGGCCAGGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCCCAGAAA 3808

Qy   1046 AspIleTyrPheLysLeuValAsnAla 1054
Db   3809 GGCATCTATTTCTCGATGGTTCAGGCT 3835
```

RESULT 8
US-09-450-105-1
; Sequence 1, Application US/09450105
; Patent No. 6169166
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: ENCODING RAT MDR1B2 AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP-50008-D1
; CURRENT APPLICATION NUMBER: US/09/450,105
; CURRENT FILING DATE: 1999-11-29
; EARLIER APPLICATION NUMBER: 09/120,513
; EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 4233
;   TYPE: DNA
;   ORGANISM: HOMO SAPIENS
US-09-450-105-1

```
Alignment Scores:
Pred. No.:              0           Length:        4233
Score:                  3086.00     Matches:       601
Percent Similarity:     74.74%      Conservative:  198
Best Local Similarity:  56.22%      Mismatches:    224
Query Match:            57.90%      Indels:        46
DB:                     4           Gaps:          7
```

US-09-873-409-4 (1-1058) x US-09-450-105-1 (1-4233)

```
Qy      1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20
Db    722 GTACTGACTTCATTTACTAATAAGGAACTCCAGGCTTATGCGAAAGCTGGAGCAGTTGCC 781

Qy     21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40
Db    782 GAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCGTTTGGAGGACAAAAGAAGGAACTT 841

Qy     41 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60
Db    842 GAA--------------------------------------------------------- 844

Qy     61 LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80
```

```
Db     845 ------------------------AGGTACAATAAAAATTTAGAAGAAGCTAAAAGAGTT 880

Qy      81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 100
Db     881 GGCATAAAGAAAGCCATCACGGCCAACATTTCCATAGGTATTGCCTACCTGTTGGTCTAT 940

Qy     101 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120
Db     941 GCGTCTTATGCACTGGCATTCTGGTATGGGACCTCCTTGGTCCTCTCAAATGAA------ 994

Qy     121 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 140
Db     995 TATTCTATTGGACAAGTGCTTACCGTCTTCTTCTCTATTTATTGGGGACTTTCAGTATT 1054

Qy     141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
Db    1055 GGACATTTAGCCCCAAACATAGAAGCCTTTGCAAATGCAAGAGGGGCAGCCTATGAAATC 1114

Qy     161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180
Db    1115 TTCAAGATAATTGATAATGAGCCAAGCATTGACAGCTTCTCAACCAAGGGACACAAACCA 1174

Qy     181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200
Db    1175 GACAGTATAATGGGAAATTTGGAATTTAAAAATGTTTACTTCAACTACCCATCACGAAGT 1234

Qy     201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
Db    1235 GAAGTTAAGATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTG 1294

Qy     221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240
Db    1295 GTTGGCAACAGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGAC 1354

Qy     241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260
Db    1355 CCCATAGAGGGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATCAATGTGAGGTAT 1414

Qy     261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280
Db    1415 CTGCGGGAAATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCCACCACGATTGCC 1474

Qy     281 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 300
Db    1475 GAAAACATTCGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAG 1534

Qy     301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320
Db    1535 GAAGCCAATGCCTATGACTTCATCATGAAACTGCCCCACAAATTTAACACCCTGGTTGGT 1594

Qy     321 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340
Db    1595 GAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTG 1654

Qy     341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360
Db    1655 GTCCGCAACCCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGC 1714

Qy     361 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380
Db    1715 GAAGCCGTGGTTCAGGCCGCTCTGGATAAGGCTAGAGAAGGCCGGACCACCATTGTGATA 1774

Qy     381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400
Db    1775 GCTCACCGCTTGTCTACAGTGCGCAATGCTGACGTCATTGCTGGTTTGATGGTGGTGTC 1834

Qy     401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420
Db    1835 ATTGTGGAGCAAGGAAATCATGAAGAGCTCATGAAAGAGAAGGGCATTTACTTCAAACTT 1894

Qy     421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440
Db    1895 GTCATGACACAG---ACTAGAGGAAATGAAATTGAACCAGGAAATAATGCTTATGAATCC 1951

Qy     441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460
Db    1952 CAAAGTGACACTGGTGCCTCTGAGTTGACTTCAGAAAAATCA---AAATCTCCTTTAATA 2008

Qy     461 AspLysAla---------------------------------GluGluSerThrGlnSer 469
Db    2009 AGGAGATCAATTCGCAGAAGTATCCACAGAAGACAAGACCAGGAGAGAAGACTTAGTTCG 2068

Qy     470 LysGlu---------IleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 486
Db    2069 AAAGAGGATGTGGATGAAGATGTGCCTATGGTTTCCTTTTGGCAGATCCTAAAGCTAAAT 2128

Qy     487 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 506
Db    2129 ATTAGTGAATGGCCCTATTTAGTTGTGGGTGTACTTTGTGCTGTTATAAATGGGTGCATA 2188

Qy     507 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGly---AsnAsnAsp 525
Db    2189 CAACCAGTGTTTGCCATAGTGTTTTCAAAGATTGTAGGGGTTTTTCAAGAGACGACGAC 2248

Qy     526 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 545
Db    2249 CATGAAACCAAACAACGGAATTGTAACTTGTTTTCCCTTCTCTTTCTGGTCATGGGAATG 2308

Qy     546 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 565
Db    2309 ATTTCTTTTGTTACGTACTTCTTTCAAGGCTTCACATTTGGCAAAGCTGGAGAGATCCTC 2368

Qy     566 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 585
Db    2369 ACCAAGCGACTCCGATACATGGTCTTCAAATCCATGCTGCGACAGGATATAAGCTGGTTT 2428

Qy     586 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 605
Db    2429 GATGACCATAAAAACACCACTGGCTCGCTGACTACCAGGCTCGCTAGTGACGCTTCTAAT 2488

Qy     606 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 625
Db    2489 GTTAAAGGGGCTATGGGCTCCAGGCTTGCTGTAGTTACCCAGAATGTAGCAAACCTTGGC 2548

Qy     626 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 645
Db    2549 ACAGGAATTATCTTATCCTTAGTCTATGGCTGGCAGCTTACACTTTTACTTGTAGTAATT 2608

Qy     646 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 665
Db    2609 ATACCACTCATTGTCTTGGGTGGAATTATTGAAATGAAACTGTTGTCTGGTCAAGCCTTG 2668

Qy     666 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 685
Db    2669 AAGGACAAGAAAGAGCTAGAGATCTCTGGGAAGATCGCTACAGAAGCAATTGAAAACTTC 2728

Qy     686 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 705
Db    2729 CGCACTGTTGTCTCTTTGACTAGGGAGCAGAAGTTTGAAACTATGTATGCCCAGAGCTTG 2788

Qy     706 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 725
Db    2789 CAGATACCATACAGAAATGCTTTGAAGAAAGCACACGTCTTTGGGATCACCTTCGCCTTC 2848

Qy     726 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 745
Db    2849 ACCCAGGCCATGATTTATTTTTCCTATGCTGCTTGTTTCCGGTTCGGTGCCTACTTGGTG 2908

Qy     746 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 765
Db    2909 GCACGAGAACTCATGACGTTTGAAAATGTTATGTTGGTATTTTCTGCTGTTGTCTTTGGT 2968

Qy     766 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 785
Db    2969 GCCATGGCAGCAGGGAATACCAGTTCATTCGCTCCTGACTACGCGAAGGCCAAAGTCTCA 3028
```

```
Qy      786 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 805
            |||:::|||:::      ::::::|||||  |||  ||||||||  |||   |||
Db     3029 GCATCCCACATCATCAGGATCATTGAGAAAATCCCCGAGATTGACAGCTACAGCACGGAG 3088

Qy      806 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 825
            |||  ||||||:::      |||||||||::::::|||        |||  |||   |||
Db     3089 GGCTTGAAGCCTAATTGGTTAGAAGGAAATGTGAAATTTAATGGAGTCATGTTCAACTAT 3148

Qy      826 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 845
            |||  ||||||::::::   :::|||:::||||||||       :::::::::|||:::
Db     3149 CCCACCCGACCCAACATCCCAGTGCTTCAGGGACTGAGCTTCGAGGTGAAGAAGGGGCAA 3208

Qy      846 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 865
            |||:::      |||||||||||||||||||||||||||   ||||||||||||:::
Db     3209 ACGCTTCGCCTGGTGGGCAGCAGTGGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTCGAG 3268

Qy      866 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 885
            |||   |||:::|||:::   |||   |||      ||||||   :::   |||:::|||
Db     3269 CGCTTCTACAACCCCATGGCTGGAACAGTGTTTCTAGATGGCAAAGAAATAAAACAACTC 3328

Qy      886 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 905
            ||||||||||||||||||:::      :::   ||||||   |||||||||:::||||||:::
Db     3329 AACGTCCAGTGGCTCCGCGCCCACCTGGGCATTGTGTCCCAGGAGCCCATCCTGTTTGAC 3388

Qy      906 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 925
            |||||||||   ||||||||||||||||||||||||||||||||||||      :::|||
Db     3389 TGCAGCATCACCGAGAACATCGCCTACGGAGACAACAGCCGTGTCGTGTCTCATGAGGAG 3448

Qy      926 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 945
            |||      ||||||   ||||||||||||   ||||||:::   ||||||||||||
Db     3449 ATCGTGAGGGCCGCCAGGGAGGCCAACATCCACCAGTTCATCGACTCACTGCCTGAGAAA 3508

Qy      946 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 965
            |||||||||:::||||||   ||||||   |||||||||||||||||||||||||||:::
Db     3509 TACAACACCAGAGTGGGAGACAAAGGGACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATC 3568

Qy      966 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 985
            ||||||||||||||||||:::::::::|||   ||||||||||||||||||||||||||||
Db     3569 GCCATCGCGCGCGCCCTCGTCAGACAGCCTCACATCTTACTTCTGGATGAAGCGACATCA 3628

Qy      986 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1005
            |||||||||   :::|||||||||||||||||||||||||||||||||||   |||
Db     3629 GCTCTGGATACGGAGAGTGAAAAGGTCGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGC 3688

Qy     1006 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1025
            |||||||||:::|||:::   ||||||||||||   ||||||||||||||||||||||||
Db     3689 CGCACCTGCATTGTGATCGCGCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTG 3748

Qy     1026 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1045
            |||:::   ||||||::::::||||||   ||||||||||||:::||||||      :::
Db     3749 GTGATTCAGAACGGCCAGGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCCCAGAAA 3808

Qy     1046 AspIleTyrPheLysLeuValAsnAla 1054
               |||||||||   :::|||   |||
Db     3809 GGCATCTATTTCTCGATGGTTCAGGCT 3835

RESULT 9
US-08-752-447-1
; Sequence 1, Application US/08752447
; Patent No. 5994088
;  GENERAL INFORMATION:
;    APPLICANT:  Mechetner, Eugene
;    APPLICANT:  Roninson, Igor B
;    TITLE OF INVENTION:  Methods and Reagents for Preparing and
;    TITLE OF INVENTION:  Using Immunoligcal Agents Specific for P-glycoprotein
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McDonnell Boehnen Hulbert & Berghoff Ltd.
;      STREET:  300 South Wacker Drive, Seventh Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/752,447
;      FILING DATE:  15-NOV-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5994088nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  95,1121
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-913-0001
;      TELEFAX:  312-913-9808
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4669 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  5'UTR
;      LOCATION:  1..424
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  425..4264
;    FEATURE:
;      NAME/KEY:  3'UTR
;      LOCATION:  4265..4669
US-08-752-447-1

Alignment Scores:
Pred. No.:              0              Length:         4669
Score:                  3073.50        Matches:        590
Percent Similarity:     74.98%         Conservative:   210
Best Local Similarity:  55.30%         Mismatches:     224
Query Match:            57.66%         Indels:         43
DB:                     2              Gaps:           6

US-09-873-409-4 (1-1058) x US-08-752-447-1 (1-4669)

Qy        4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
            |||   |||   |||||||||   ||||||:::|||||||||||||||||||||||||||
Db     1136 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 1195

Qy       24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
            |||::::::||||||||||||||||||||||||      |||:::|||||||||:::
Db     1196 TTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAAAAGAAAGAACTTGAA------ 1249

Qy       44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63

Db     1249 ------------------------------------------------------------ 1249

Qy       64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
                           ||||||   :::||||||::::::||||||      |||||||||
Db     1250 ---------------AGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1294

Qy       84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
            :::   |||   :::   :::|||:::||||||   :::      :::      :::|||
Db     1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1354

Qy      104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
               :::||||||||||||||||||:::|||:::|||:::||||||      |||:::|||
Db     1355 GCTGTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1408

Qy      124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
            |||   ||||||   |||||||||||||||:::      ::::::   :::|||   |||
Db     1409 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1468

Qy      144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
```

```
Db   1469 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528

Qy    164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db   1529 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1588

Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db   1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1648

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db   1649 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1708

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db   1709 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1768

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db   1769 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1828

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db   1829 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1888

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db   1889 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1948

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db   1949 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2008

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db   2009 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2068

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db   2069 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2128

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db   2129 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 2188

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db   2189 TTGCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 2248

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   2249 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 2308

Qy    424 Gln------------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
Db   2309 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 2368

Qy    438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 452
Db   2369 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2428

Qy    453 ---LysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLys 470
Db   2429 CGTAGGAGTGTCCGTGGATCACAAGCCCAACACAGAAAGCTTAGTACCAAAGAGGCTCTG 2488

Qy    471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
Db   2489 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2548

Qy    491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db   2549 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2608

Qy    511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
Db   2609 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2668

Qy    530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db   2669 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2728

Qy    550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db   2729 ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2788

Qy    570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
Db   2789 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTCATGACCCTAAA 2848

Qy    590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db   2849 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2908

Qy    610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db   2909 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2968

Qy    630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db   2969 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 3028

Qy    650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db   3029 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTTGCTGGACAAGCACTGAAAGATAAGAAA 3088

Qy    670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db   3089 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 3148

Qy    690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db   3149 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 3208

Qy    710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db   3209 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 3268

Qy    730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db   3269 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 3328

Qy    750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db   3329 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 3388

Qy    770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db   3389 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3448

Qy    790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db   3449 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3508

Qy    810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db   3509 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3568

Qy    830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db   3569 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3628

Qy    850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db   3629 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3688
```

```
Qy     870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
           |||:::   |||:::||||||  ||||||   :::  |||  ||||||||||||||
Db    3689 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3748

Qy     890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
           ||||||:::   :::   ||||||   |||||||||:::||||||:::||||||||||||
Db    3749 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3808

Qy     910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
           ||||||||||||||||||||||||||||||||||||       :::||||||     |||
Db    3809 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3868

Qy     930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
           |||   ||||||||||||:::|||||||||   ||||||   ||||||:::|||:::
Db    3869 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3928

Qy     950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
           ||||||   ||||||   ||||||||||||||||||||||||:::||||||||||||
Db    3929 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3988

Qy     970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
              |||:::::::::|||   ||||||||||||||||||||||||||||||||||||
Db    3989 CGCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 4048

Qy     990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
           :::||||||||||||||||||   ||||||||||||||||||   ||||||||||||:::
Db    4049 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 4108

Qy    1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
           |||:::   ||||||||||||   ||||||||||||||||||||||||||||    |||
Db    4109 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 4168

Qy    1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
           |||::::::||||||   ||||||||||||:::||||||       :::   ||||||||
Db    4169 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 4228

Qy    1050 LysLeuValAsnAlaGlnSer 1056
              :::|||:::   |||:::
Db    4229 TCAATGGTCAGTGTCCAGGCT 4249
```

RESULT 10
US-09-316-167-1
; Sequence 1, Application US/09316167
; Patent No. 6365357
; GENERAL INFORMATION:
; APPLICANT: Mechetner, Eugene
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods and Reagents for Preparing and
; TITLE OF INVENTION: Using Immunoligcal Agents Specific for P-glycoprotein
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff Ltd.
; STREET: 300 South Wacker Drive, Seventh Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/316,167
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/752,447
; FILING DATE: 15-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6365357nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 95,1121
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-9808
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..424
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 425..4264
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 4265..4669
US-09-316-167-1

```
Alignment Scores:
Pred. No.:              0           Length:        4669
Score:                  3073.50     Matches:       590
Percent Similarity:     74.98%      Conservative:  210
Best Local Similarity:  55.30%      Mismatches:    224
Query Match:            57.66%      Indels:        43
DB:                     4           Gaps:          6
```

US-09-873-409-4 (1-1058) x US-09-316-167-1 (1-4669)

```
Qy       4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
           |||   |||   |||||||||   ||||||:::||||||||||||||||||||||||||||
Db    1136 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 1195

Qy      24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
           |||::::::||||||||||||||||||||||||   |||:::||||||||||||:::
Db    1196 TTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAAAAGAAAGAACTTGAA------ 1249

Qy      44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63

Db    1249 ------------------------------------------------------------ 1249

Qy      64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
                            ||||||   :::||||||::::::||||||     ||||||||
Db    1250 ---------------AGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1294

Qy      84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
           :::   |||   :::   :::|||:::||||||   :::       :::      :::|||
Db    1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1354

Qy     104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
           :::|||||||||||||||||||:::|||:::|||:::||||||      |||:::|||
Db    1355 GCTGTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1408

Qy     124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
           |||   ||||||   ||||||||||||||||:::      ::::::   :::|||   |||
Db    1409 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1468

Qy     144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
              |||      |||   ||||||   ||||||||||||||:::   ||||||::::::
Db    1469 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528

Qy     164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
           ||||||   ||||||||||||||::::::|||   :::|||:::||||||::::::|||
Db    1529 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1588

Qy     184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
           :::|||   :::||||||:::||||||   |||:::||||||||||       :::|||
Db    1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1648

Qy     204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
```

```
Db   1649 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1708

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db   1709 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1768

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db   1769 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1828

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db   1829 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1888

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db   1889 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1948

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db   1949 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2008

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db   2009 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2068

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db   2069 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2128

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db   2129 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 2188

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db   2189 TTGCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 2248

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   2249 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 2308

Qy    424 Gln------------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
Db   2309 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 2368

Qy    438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 452
Db   2369 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2428

Qy    453 ---LysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLys 470
Db   2429 CGTAGGAGTGTCCGTGGATCACAAGCCCAACACAGAAAGCTTAGTACCAAAGAGGCTCTG 2488

Qy    471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
Db   2489 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2548

Qy    491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db   2549 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2608

Qy    511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
Db   2609 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2668

Qy    530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db   2669 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2728

Qy    550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db   2729 ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2788

Qy    570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
Db   2789 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTCATGACCCTAAA 2848

Qy    590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db   2849 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2908

Qy    610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db   2909 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2968

Qy    630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db   2969 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 3028

Qy    650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db   3029 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTTGCTGGACAAGCACTGAAAGATAAGAAA 3088

Qy    670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db   3089 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 3148

Qy    690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db   3149 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 3208

Qy    710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db   3209 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 3268

Qy    730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db   3269 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 3328

Qy    750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db   3329 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 3388

Qy    770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db   3389 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3448

Qy    790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db   3449 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3508

Qy    810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db   3509 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3568

Qy    830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db   3569 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3628

Qy    850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db   3629 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3688

Qy    870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
Db   3689 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3748

Qy    890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db   3749 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3808

Qy    910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db   3809 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3868
```

```
Qy     930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
           |||      ||||||||||:::||||||||   ||||||     |||||:::|||:::
Db    3869 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3928

Qy     950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
           ||||||   ||||||   ||||||||||||||||||||||||||:::||||||||||||
Db    3929 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3988

Qy     970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
              |||:::::::::|||   ||||||||||||||||||||||||||||||||||||
Db    3989 CGCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 4048

Qy     990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
           :::||||||||||||||||||   |||||||||||||||||   ||||||||||||:::
Db    4049 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 4108

Qy    1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
           |||:::   ||||||||||||   |||||||||||||||||||||||||||      |||
Db    4109 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 4168

Qy    1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
           |||::::::||||||   ||||||||||||:::||||||      :::   |||||||||
Db    4169 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 4228

Qy    1050 LysLeuValAsnAlaGlnSer 1056
              :::|||:::   |||:::
Db    4229 TCAATGGTCAGTGTCCAGGCT 4249
```

RESULT 11
US-08-583-276-18
; Sequence 18, Application US/08583276
; Patent No. 5837536
; GENERAL INFORMATION:
;   APPLICANT: McDonagh, Kevin T.
;   APPLICANT: Nienhuis, Arthur
;   APPLICANT: Tolstoshev, Paul
;   TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN
;   TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED
;   TITLE OF INVENTION: SELECTION OF CELLS TRANSDUCED WITH SUCH GENES
;   NUMBER OF SEQUENCES: 19
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
;   ADDRESSEE: Cecchi & Stewart
;   STREET: 6 Becker Farm Road
;   CITY: Roseland
;   STATE: New Jersey
;   COUNTRY: USA
;   ZIP: 07068
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3.5 inch diskette
;   COMPUTER: IBM PS/2
;   OPERATING SYSTEM: PC-DOS
;   SOFTWARE: DW4.V2
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/583,276
;   FILING DATE: 05-JAN-1996
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/332,444
;   FILING DATE: 31-OCT-1994
;   APPLICATION NUMBER: 07/887,712
;   FILING DATE: 22-MAY-1992
; INFORMATION FOR SEQ ID NO: 18:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4669 bases
;   TYPE: nucleic acid
;   STRANDEDNESS: singular
;   TOPOLOGY: linear
;   MOLECULE TYPE:
;   DESCRIPTION: Genomic DNA
US-08-583-276-18

```
Alignment Scores:
Pred. No.:               0              Length:         4669
Score:                   3056.50        Matches:        588
Percent Similarity:      74.86%         Conservative:   210
Best Local Similarity:   55.16%         Mismatches:     225
Query Match:             57.35%         Indels:         43
DB:                      2              Gaps:           6

US-09-873-409-4 (1-1058) x US-08-583-276-18 (1-4669)

Qy       4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
           |||   |||   ||||||||   ||||||:::||||||||||||||||||||||||||||
Db    1136 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 1195

Qy      24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
           |||::::::|||||||||||||||||||||||||      |||:::||||||||||:::
Db    1196 TTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAAAAGAAAGAACTTGAA------ 1249

Qy      44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63

Db    1249 ------------------------------------------------------------ 1249

Qy      64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
                          ||||||   :::||||||:::::::||||||      |||||||||
Db    1250 ---------------AGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1294

Qy      84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
           :::   |||   :::   :::|||:::||||||   :::         :::      :::|||
Db    1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1354

Qy     104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
              ||||||||||||||||||||||||:::|||:::|||:::|||||||      |||:::|||
Db    1355 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1408

Qy     124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
           |||   ||||||   ||||||||||||||||:::         ::::::   :::|||   |||
Db    1409 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1468

Qy     144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
              |||      |||   ||||||   ||||||||||||||:::   ||||||::::::
Db    1469 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528

Qy     164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
           ||||||   ||||||||||||||:::::::|||   :::|||:::||||||::::::|||
Db    1529 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1588

Qy     184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
           :::|||   :::||||||:::||||||   |||:::|||||||||||      :::|||
Db    1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1648

Qy     204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
           ||||||||||||||||||:::::::::||||||:::||||||   |||||||||
Db    1649 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTCCCCCTGGTTGGAAAC 1708

Qy     224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
           :::|||   ||||||||||||   |||||||||:::||||||||||||||||||   :::
Db    1709 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1768

Qy     244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
           |||   :::   |||         |||||||||   :::|||||||||      |||:::
Db    1769 GGGATGGTCAGTGTTCATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1828

Qy     264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
              ||||||||||||||||||||||||||||||||||   ||||||||||:::   ||||||
Db    1829 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1888

Qy     284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
           :::|||||||||::::::||||||   :::|||:::|||:::|||   :::|||||||||
Db    1889 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1948

Qy     304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
           ||||||||||||||||||:::   |||:::||||||:::||||||||||||||||:::|||
Db    1949 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2008
```

```
Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db   2009 CCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2068

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db   2069 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2128

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db   2129 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 2188

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db   2189 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 2248

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   2249 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 2308

Qy    424 Gln------------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
Db   2309 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 2368

Qy    438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 452
Db   2369 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCACGTCTAATAAGAAAAAGATCAACT 2428

Qy    453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470
Db   2429 CGTAGGAGTGTCCGTGGATCACAACCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2488

Qy    471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
Db   2489 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2548

Qy    491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db   2549 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2608

Qy    511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
Db   2609 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2668

Qy    530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db   2669 CGACAGAATACTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2728

Qy    550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db   2729 ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2788

Qy    570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
Db   2789 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2848

Qy    590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db   2849 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2908

Qy    610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db   2909 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAATTCTTGGGACAGGAATAATT 2968

Qy    630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db   2969 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 3028

Qy    650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db   3029 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 3088

Qy    670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db   3089 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 3148

Qy    690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db   3149 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 3208

Qy    710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db   3209 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 3268

Qy    730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db   3269 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTA 3328

Qy    750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db   3329 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 3388

Qy    770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db   3389 GGGCAAGTGAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3448

Qy    790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db   3449 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3508

Qy    810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db   3509 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3568

Qy    830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db   3569 GACATCCCAGTCCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3628

Qy    850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db   3629 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3688

Qy    870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
Db   3689 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3748

Qy    890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db   3749 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3808

Qy    910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db   3809 GAGAACATTGCCTATGGACACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3868

Qy    930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db   3869 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3928

Qy    950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db   3929 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3988

Qy    970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db   3989 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 4048

Qy    990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db   4049 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAACCCAGAGAAGGCCGCACCTGCATT 4108

Qy   1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db   4109 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 4168

Qy   1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
```

```
            |||:::::|||:::  |||||||||||:::|||||       :::   ||||||||
Db   4169 GGCAGAGTCAAGCAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 4228

Qy   1050 LysLeuValAsnAlaGln 1055
              :::|||:::  |||
Db   4229 TCAATGGTCAGTGTCCAG 4246

RESULT 12
US-08-461-823-1
; Sequence 1, Application US/08461823
; Patent No. 5593840
;  GENERAL INFORMATION:
;    APPLICANT:  Bhatnagar, Satish K.
;    APPLICANT:  George Jr., Albert L.
;    APPLICANT:  Nazarenko, Irina
;    TITLE OF INVENTION:  AMPLIFICATION OF NUCLEIC ACID SEQUENCES
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OncorPharm, Inc.
;      STREET:  200 Perry Parkway
;      CITY:  Gaithersburg
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20877
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/461,823
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/168,621
;      FILING DATE:  16-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/010,433
;      FILING DATE:  27-JAN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Karta, Glenn E.
;      REGISTRATION NUMBER:  30,649
;      REFERENCE/DOCKET NUMBER:  PA-0012 CIP 2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  301 527-2058
;      TELEFAX:  301 208-6997
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  2726 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-461-823-1

Alignment Scores:
Pred. No.:              3.5e-246     Length:         2726
Score:                  2280.00      Matches:        435
Percent Similarity:     76.55%       Conservative:   159
Best Local Similarity:  56.06%       Mismatches:     168
Query Match:            42.78%       Indels:         14
DB:                     1            Gaps:           4

US-09-873-409-4 (1-1058) x US-08-461-823-1 (1-2726)

Qy    295 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 314
          :::|||:::|||   :::||||||||||||||||||||||||||:::   |||:::|||
Db      2 ATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAA 61

Qy    315 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 334
          |||:::|||||||||||||||:::||||||||:::||||||||||||||||||||||||
Db     62 TTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATC 121

Qy    335 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 354
          ||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db    122 GCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCA 181

Qy    355 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 374
          ||||||||:::||||||::::::    ||||||   ||||||:::||||||   ||||||
Db    182 GCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGT 241

Qy    375 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 394
          |||||||||||||||:::||||||||||||||||||:::|||:::||||||:::|||
Db    242 CGGACCACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCT 301

Qy    395 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 414
                   ||||||::::::   |||||||||   |||   |||||||||   ::::::
Db    302 GGTTTCGATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAA 361

Qy    415 GlyLeuTyrTyrSerLeuValMetSerGln-----------------AspIleLysLys 428
          |||:::|||:::   ||||||      |||                  :::::::::
Db    362 GGCATTTACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAAT 421

Qy    429 AlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPro 448
          |||   ::::::   :::|||       :::   |||       :::|||
Db    422 GCAGCTGATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGA 481

Qy    449 LeuHisSerVal-----------------LysSerIleLys---SerAspPheIleAsp 461
                :::               :::|||::::::   |||         |||
Db    482 TCCAGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGAC 541

Qy    462 LysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLys 481
          :::         :::   ::::::   :::   |||:::|||   ||||||      :::
Db    542 AGAAAGCTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGG 601

Qy    482 IleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerVal 501
          |||:::|||||||||   |||||||||:::   |||:::|||         ::::::
Db    602 ATTATGAAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATT 661

Qy    502 LeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPhe 521
          :::||||||   :::   |||   |||:::|||||||||:::|||||||||   :::|||
Db    662 ATAAATGGAGGCCTGCAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTT 721

Qy    522 GlyAsn---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 540
                :::|||      |||   :::   ::::::   ::::::|||::::::|||
Db    722 ACAAGAATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTT 781

Qy    541 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 560
          :::   ||||||:::|||   |||:::::::::|||:::||||||      :::|||:::
Db    782 CTAGCCCTTGGAATTATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAA 841

Qy    561 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 580
          ||||||||||||||||||   |||||||||::::::   |||::::::||||||   |||
Db    842 GCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAG 901

Qy    581 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600
          |||::::::|||||||||:::   :::|||:::||||||   |||||||||   ||||||
Db    902 GATGTGAGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCC 961

Qy    601 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620
             |||   ||||||::::::||||||   ||||||||||||   |||:::|||||||||
Db    962 AATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAAT 1021

Qy    621 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 640
                   |||:::|||      :::||||||||||||||||||||||||::::::|||
Db   1022 ATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTG 1081

Qy    641 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 660
          |||:::|||:::|||   |||::::::|||:::   |||::::::|||         :::
Db   1082 TTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTG 1141

Qy    661 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 680
          :::|||   |||   |||||||||:::||||||:::   ||||||||||||||||||
Db   1142 TCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAA 1201
```

```
Qy    681 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 700
          |||:::||||||  ||||||:::|||||||||||:::|||:::  ||||||  |||
Db   1202 GCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATG 1261

Qy    701 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720
          |||  :::  ||||||    :::||||||:::   :::||||||  |||  |||
Db   1262 TATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGA 1321

Qy    721 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 740
                ::::::|||:::  |||  :::||||||:::||||||    |||||||||
Db   1322 ATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTT 1381

Qy    741 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760
          ||||||||||||:::          |||:::  |||   :::   :::||||||:::
Db   1382 GGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCA 1441

Qy    761 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 780
          |||:::   :::||||||||||||:::|||:::            ||||||:::|||:::
Db   1442 GCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCC 1501

Qy    781 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 800
          |||||||||   |||||||||:::      ::::::||||||  |||   ||||||
Db   1502 AAAGCCAAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGAC 1561

Qy    801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820
          |||  |||  ||||||     |||:::|||  |||||||||:::  |||  |||
Db   1562 AGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAA 1621

Qy    821 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 840
          |||  |||   ||||||  |||||||||:::  :::|||:::||||||||||||
Db   1622 GTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAG 1681

Qy    841 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 860
          ::::::::|||:::|||:::|||  ||||||||||||||||||||||||||||
Db   1682 GTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTG 1741

Qy    861 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 880
          ||||||||||||:::|||  ||||||||:::   |||:::||||||  ||||||
Db   1742 GTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAA 1801

Qy    881 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 900
          :::  |||  ||||||||||||||||||||||:::   :::  ||||||  ||||||
Db   1802 GAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAG 1861

Qy    901 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 920
          |||:::||||||:::||||||||||||||||||||||||||||||||||||||||||
Db   1862 CCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTG 1921

Qy    921 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 940
          |||     :::||||||   ||||||  |||||||||||:::|||||||||
Db   1922 GTGTCACAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAG 1981

Qy    941 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 960
                ||||||  ||||||:::|||:::||||||  ||||||  ||||||||||||||
Db   1982 TCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGC 2041

Qy    961 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 980
          ||||||||||||:::|||||||||||||||||||:::::::::|||  ||||||||||||
Db   2042 CAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTG 2101

Qy    981 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1000
          ||||||||||||||||||||||||   :::||||||||||||||||||   |||||||||
Db   2102 GATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGAC 2161

Qy   1001 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1020
          |||||||||  ||||||||||||:::|||:::   |||||||||||   |||||||||
Db   2162 AAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAAT 2221

Qy   1021 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1040
          ||||||||||||||||||   ||||||:::::||||||  ||||||||||||:::
Db   2222 GCAGACTTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAG 2281

Qy   1041 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1056
          ||||||     :::  |||||||||   :::|||:::  |||:::
Db   2282 CTGCTGGCACAGAAGGCATCTATTTTCAATGGTCAGTGTCCAGGCT 2329
```

RESULT 13
US-08-996-545-1
; Sequence 1, Application US/08996545
; Patent No. 5928898
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atrD of
; TITLE OF INVENTION: Aspergillus Nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,545
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4002
US-08-996-545-1

Alignment Scores:

Pred. No.:	6.76e-222	Length:	4002
Score:	2067.00	Matches:	447
Percent Similarity:	59.43%	Conservative:	199
Best Local Similarity:	41.12%	Mismatches:	381
Query Match:	38.78%	Indels:	60
DB:	2	Gaps:	9

US-09-873-409-4 (1-1058) x US-08-996-545-1 (1-4002)

```
Qy      2 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 21
          :::|||    :::  |||  |||  :::|||          |||  |||||||||
Db    817 ATCATCAAGTACAGCAAAAAGTCGCTTGACAGCTACGGTGCAGGCGGCACTGTTGCGGAA 876

Qy     22 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 41
          ||||||:::||||||||||||         ||||||   |||:::|||     :::
Db    877 GAGGTCATCAGCTCCATCAGAAATGCCACAGCGTTTGGCACCCAAGACAAGCTTGCGAAG 936

Qy     42 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeu 61
```

```
Db    936 ------------------------------------------------------------ 936
Qy     62 SerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 81
Db    937 ---------------------CAGTATGAGGTCCACTTAGACGAAGCTGAGAAATGGGGA 975
Qy     82 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 101
Db    976 ACAAAGAACCAGATTGTCATGGGTTTCATGATTGGCGCCATGTTTGGCCTTATGTACTCG 1035
Qy    102 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 121
Db   1036 AACTACGGTCTTGGCTTCTGGATGGGTTCTCGTTTCCTGGTAGATGGT------GCAGTC 1089
Qy    122 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 141
Db   1090 GATGTGGGTGATATTCTCACAGTTCTCATGGCCATCTTGATCGGATCGTTCTCCTTGGGG 1149
Qy    142 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 161
Db   1150 AACGTTAGTCCAAATGCTCAAGCATTTACAAACGCTGTGGCCGCGGCCGCAAAGATATTT 1209
Qy    162 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 181
Db   1210 GGAACGATCGATCGCCAGTCCCCATTAGATCCATATTCGAACGAAGGGAAGACGCTCGAC 1269
Qy    182 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 201
Db   1270 CATTTTGAGGGCCACATTGAGTTACGCAATGTCAAGCATATTTACCCATCTAGACCCGAG 1329
Qy    202 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 221
Db   1330 GTCACCGTCATGGAGGATGTTTCTCTGTCAATGCCCGCTGGAAAAACAACCGCTTTAGTC 1389
Qy    222 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 241
Db   1390 GGCCCCTCTGGCTCTGGAAAAAGTACGGTGGTCGGCTTGGTTGAGCGATTCTACATGCCT 1449
Qy    242 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 261
Db   1450 GTTCGCGGTACGGTTTTGCTGGATGGCCATGACATCAAGGACCTCAATCTCCGCTGGCTT 1509
Qy    262 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 281
Db   1510 CGCCAACAGATCTCTTTGGTTAGCCAGGAGCCTGTTCTTTTTGGCACGACGATTTATAAG 1569
Qy    282 AsnIleLysTyrGly--------------------------ArgAspAspValThrAsp 292
Db   1570 AATATTAGGCACGGTCTCATCGGCACAAAGTACGAGAATGAATCCGAGGATAAGGTCCGG 1629
Qy    293 GluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhePro 312
Db   1630 GAACTCATCGAGAACGCGGCAAAAATGGCGAATGCTCATGACTTTATTACTGCCTTGCCT 1689
Qy    313 AsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGln 332
Db   1690 GAAGGTTATGAGACCAATGTTGGGCAGCGTGGCTTTCTCCTTTCAGGTGGCCAGAAACAG 1749
Qy    333 ArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAla 352
Db   1750 CGCATTGCAATCGCCCGTGCCGTTGTTAGTGACCCAAAAATCCTGCTCCTGGATGAAGCT 1809
Qy    353 ThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSer 372
Db   1810 ACTTCGGCCTTGGACACAAAATCCGAAGGCGTGGTTCAAGCAGCTTTGGAGAGGGCAGCT 1869
Qy    373 LysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeu 392
Db   1870 GAAGGCCGAACTACTATTGTGATCGCTCATCGCCTTTCCACGATCAAAACGGCGCACAAC 1929
Qy    393 IleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAla 412
Db   1930 ATTGTGGTTCTGGTCAATGGCAAAATTGCTGAACAAGGAACTCACGATGAATTGGTTGAC 1989
Qy    413 LysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGln 432
Db   1990 CGCGGAGGCGCTTATCGCAAACTTGTGGAGGCTCAACGTATCAATGAACAGAAGGAAGCT 2049
Qy    433 MetGluSerMetThrTyrSerThrGluArgLysThrAsnSer------LeuProLeuHis 450
Db   2050 GACGCCTTGGAGGACGCCGACGCTGAGGATCTCACGAATGCAGATATTGCCAAAATCAAA 2109
Qy    451 SerValLysSerIleLysSerAspPhe------------------IleAspLys------ 462
Db   2110 ACTGCGTCAAGCGCATCATCCGATCTCGACGGAAAACCCACAACCATTGACCGCACGGGC 2169
Qy    463 AlaGluGluSerThrGlnSerLysGluIleSer------------------LeuProGlu 476
Db   2170 ACCCACAAGTCTGTTTCCAGCGCGATTCTTTCTAAAAGACCCCCCGAAACAACTCCGAAA 2229
Qy    477 ValSerLeuLeuLysIleLeuLys------------LeuAsnLysProGluTrpProPhe 492
Db   2230 TACTCATTATGGACGCTGCTCAAATTTGTTGCTTCCTTCAACCGCCCTGAAATCCCGTAC 2289
Qy    493 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 512
Db   2290 ATGCTCATCGGTCTTGTCTTCTCAGTGTTAGCTGGTGGTGGCCAACCCACGCAAGCAGTG 2349
Qy    513 IlePheAlaLysIleIleThrMetPheGly------AsnAsnAspLysThrThrLeuLys 530
Db   2350 CTATATGCTAAAGCCATCAGCACACTCTCGCTCCCAGAATCACAATATAGCAAGCTTCGA 2409
Qy    531 HisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSer 550
Db   2410 CATGATGCGGATTTCTGGTCATTGATGTTCTTCGTGGTTGGTATCATTCAGTTTATCACG 2469
Qy    551 TyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArg 570
Db   2470 CAGTCAACCAATGGTGCTGCATTTGCCGTATGCTCCGAGAGACTTATTCGTCGCGCGAGA 2529
Qy    571 HisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsn 590
Db   2530 AGCACTGCCTTTCGGACGATACTCCGTCAAGACATTGCTTTCTTTGACAAGGAAGAGAAT 2589
Qy    591 SerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThr 610
Db   2590 AGCACCGGCGCTCTGACCTCTTTCCTGTCCACCGAGACGAAGCATCTCTCCGGTGTTAGC 2649
Qy    611 GlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIle 630
Db   2650 GGTGTGACTCTAGGCACGATCTTGATGACCTCCACGACCCTAGGAGCGGCTATCATTATT 2709
Qy    631 SerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAla 650
Db   2710 GCCCTGGCGATTGGGTGGAAATTGGCCTTAGTTTGTATCTCGGTTGTGCCGGTTCTCCTG 2769
Qy    651 ValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGlu 670
Db   2770 GCATGCGGTTTCTACCGATTCTATATGCTAGCCCAGTTTCAATCACGCTCCAAGCTTGCT 2829
Qy    671 LeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSer 690
Db   2830 TATGAGGGATCTGCAAACTTTGCTTGCGAGGCTACATCGTCTATCCGCACAGTTGCGTCA 2889
Qy    691 LeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArg 710
Db   2890 TTAACCCGGGAAAGGGATGTCTGGGAGATTTACCATGCCCAGCTTGACGCACAAGGCAGG 2949
Qy    711 AsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIle 730
Db   2950 ACCAGTCTAATCTCTGTCTTGAGGTCATCCCTGTTATATGCGTCGTCGCAGGCACTTGTT 3009
Qy    731 TyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMet 750
Db   3010 TTCTTCTGCGTTGCGCTCGGGTTTTGGTACGGAGGGACACTTCTTGGTCACCACGAGTAT 3069
```

```
Qy     751 ThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGly 770
Db    3070 GACATTTTCCGCTTCTTTGTTTGTTTCTCCGAGATTCTCTTTGGTGCTCAATCCGCGGGC 3129

Qy     771 LysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPhe 790
Db    3130 ACCGTCTTTTCCTTTGCACCAGACATGGGCAAGGCGAAGAATGCGGCCGCCGAATTCCGA 3189

Qy     791 AlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAsp 810
Db    3190 CGACTGTTCGACCGAAAGCCACAAATTGATAACTGGTCTGAAGAGGGCGAGAAGCTCGAA 3249

Qy     811 ThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAsp 830
Db    3250 ACGGTGGAAGGTGAAATCGAATTTAGGAACGTGCACTTCAGATACCCGACCCGCCCAGAA 3309

Qy     831 ValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheVal 850
Db    3310 CAGCCTGTCCTGCGCGGCTTGGACCTGACCGTGAAGCCTGGACAATATGTTGCGCTTGTC 3369

Qy     851 GlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspPro 870
Db    3370 GGACCCAGCGGTTGTGGCAAGAGTACCACCATTGCATTGCTTGAGCGCTTTTACGATGCG 3429

Qy     871 ValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeu 890
Db    3430 ATTGCCGGGTCCATCCTTGTTGATGGGAAGGACATAAGTAAACTAAATATCAACTCCTAC 3489

Qy     891 ArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGlu 910
Db    3490 CGCAGCTTTCTGTCACTGGTCAGCCAGGAGCCGACACTGTACCAGGGCACCATCAAGGAA 3549

Qy     911 AsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAla 930
Db    3550 AACATCTTACTTGGTATTGTCGAAGATGACGTACCGGAAGAATTCTTGATTAAGGCTTGC 3609

Qy     931 AsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnVal 950
Db    3610 AAGGACGCTAATATCTACGACTTCATCATGTCGCTCCCGGAGGGCTTTAATACAGTTGTT 3669

Qy     951 GlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAla 970
Db    3670 GGCAGCAAGGGAGGCATGTTGTCTGGCGGCCAAAAGCAACGTGTGGCCATTGCCCGAGCC 3729

Qy     971 LeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAsp 990
Db    3730 CTTCTTCGGGATCCCAAAATCCTTCTTCTCGATGAAGCGACGTCAGCCCTCGACTCCGAG 3789

Qy     991 SerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuVal 1010
Db    3790 TCAGAAAAGGTCGTCCAGGCGGCTTTGGATGCCGCTGCCCGAGGCCGAACCACAATCGCC 3849

Qy    1011 ValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGly 1030
Db    3850 GTTGCACACCGACTCAGCACGATTCAAAAGGCGGACGTTATCTATGTTTTCGACCAAGGC 3909

Qy    1031 LysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLys 1050
Db    3910 AAGATCGTCGAAAGCGGAACGCACAGCGAACTGGTCCAGAAAAAGGGCCGGTACTACGAG 3969

Qy    1051 LeuValAsnAlaGlnSerVal 1057
Db    3970 CTGGTCAACTTGCAGAGCTTG 3990
```

RESULT 14
US-08-996-545-3
; Sequence 3, Application US/08996545
; Patent No. 5928898
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atrD of
; TITLE OF INVENTION: Aspergillus Nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,545
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: mRNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-996-545-3

Alignment Scores:

Pred. No.:	6.76e-222	Length:	4002
Score:	2067.00	Matches:	447
Percent Similarity:	59.43%	Conservative:	199
Best Local Similarity:	41.12%	Mismatches:	381
Query Match:	38.78%	Indels:	60
DB:	2	Gaps:	9

US-09-873-409-4 (1-1058) x US-08-996-545-3 (1-4002)

```
Qy       2 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 21
Db     817 AUCAUCAAGUACAGCAAAAAGUCGCUUGACAGCUACGGUGCAGGCGGCACUGUUGCGGAA 876

Qy      22 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 41
Db     877 GAGGUCAUCAGCUCCAUCAGAAAUGCCACAGCGUUUGGCACCCAAGACAAGCUUGCGAAG 936

Qy      42 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeu 61
Db     936 ------------------------------------------------------------ 936

Qy      62 SerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 81
Db     937 --------------------CAGUAUGAGGUCCACUUAGACGAAGCUGAGAAAUGGGGA 975

Qy      82 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 101
Db     976 ACAAAGAACCAGAUUGUCAUGGGUUUCAUGAUUGGCGCCAUGUUUGGCCUUAUGUACUCG 1035

Qy     102 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 121
Db    1036 AACUACGGUCUUGGCUUCUGGAUGGGUUCUCGUUUCCUGGUAGAUGGU------GCAGUC 1089

Qy     122 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 141
```

```
Db   1090 GAUGUGGGUGAUAUUCUCACAGUUCUCAUGGCCAUCUUGAUCGGAUCGUUCUCCUUGGGG 1149

Qy    142 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 161
Db   1150 AACGUUAGUCCAAAUGCUCAAGCAUUUACAAACGCUGUGGCCGCGGCCGCAAAGAUAUUU 1209

Qy    162 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 181
Db   1210 GGAACGAUCGAUCGCCAGUCCCCAUUAGAUCCAUAUUCGAACGAAGGGAAGACGCUCGAC 1269

Qy    182 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 201
Db   1270 CAUUUUGAGGGCCACAUUGAGUUACGCAAUGUCAAGCAUAUUUACCCAUCUAGACCCGAG 1329

Qy    202 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 221
Db   1330 GUCACCGUCAUGGAGGAUGUUUCUCUGUCAAUGCCCGCUGGAAAAACAACCGCUUUAGUC 1389

Qy    222 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 241
Db   1390 GGCCCCUCUGGCUCUGGAAAAAGUACGGUGGUCGGCUUGGUUGAGCGAUUCUACAUGCCU 1449

Qy    242 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 261
Db   1450 GUUCGCGGUACGGUUUUGCUGGAUGGCCAUGACAUCAAGGACCUCAAUCUCCGCUGGCUU 1509

Qy    262 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 281
Db   1510 CGCCAACAGAUCUCUUUGGUUAGCCAGGAGCCUGUUCUUUUUGGCACGACGAUUUAUAAG 1569

Qy    282 AsnIleLysTyrGly--------------------------ArgAspAspValThrAsp 292
Db   1570 AAUAUUAGGCACGGUCUCAUCGGCACAAAGUACGAGAAUGAAUCCGAGGAUAAGGUCCGG 1629

Qy    293 GluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhePro 312
Db   1630 GAACUCAUCGAGAACGCGGCAAAAAUGGCGAAUGCUCAUGACUUUAUUACUGCCUUGCCU 1689

Qy    313 AsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGln 332
Db   1690 GAAGGUUAUGAGACCAAUGUUGGGCAGCGUGGCUUUCUCCUUUCAGGUGGCCAGAAACAG 1749

Qy    333 ArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAla 352
Db   1750 CGCAUUGCAAUCGCCCGUGCCGUUGUUAGUGACCCAAAAAUCCUGCUCCUGGAUGAAGCU 1809

Qy    353 ThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSer 372
Db   1810 ACUUCGGCCUUGGACACAAAAUCCGAAGGCGUGGUUCAAGCAGCUUUGGAGAGGGCAGCU 1869

Qy    373 LysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeu 392
Db   1870 GAAGGCCGAACUACUAUUGUGAUCGCUCAUCGCCUUUCCACGAUCAAAACGGCGCACAAC 1929

Qy    393 IleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAla 412
Db   1930 AUUGUGGUUCUGGUCAAUGGCAAAAUUGCUGAACAAGGAACUCACGAUGAAUUGGUUGAC 1989

Qy    413 LysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGln 432
Db   1990 CGCGGAGGCGCUUAUCGCAAACUUGUGGAGGCUCAACGUAUCAAUGAACAGAAGGAAGCU 2049

Qy    433 MetGluSerMetThrTyrSerThrGluArgLysThrAsnSer------LeuProLeuHis 450
Db   2050 GACGCCUUGGAGGACGCCGACGCUGAGGAUCUCACGAAUGCAGAUAUUGCCAAAAUCAAA 2109

Qy    451 SerValLysSerIleLysSerAspPhe------------------IleAspLys------ 462
Db   2110 ACUGCGUCAAGCGCAUCAUCCGAUCUCGACGGAAAACCCACAACCAUUGACCGCACGGGC 2169

Qy    463 AlaGluGluSerThrGlnSerLysGluIleSer------------------LeuProGlu 476
Db   2170 ACCCACAAGUCUGUUUCCAGCGCGAUUCUUUCUAAAAGACCCCCCGAAACAACUCCGAAA 2229

Qy    477 ValSerLeuLeuLysIleLeuLys------------LeuAsnLysProGluTrpProPhe 492
Db   2230 UACUCAUUAUGGACGCUGCUCAAAUUUGUUGCUUCCUUCAACCGCCCUGAAAUCCCGUAC 2289

Qy    493 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 512
Db   2290 AUGCUCAUCGGUCUUGUCUUCUCAGUGUUAGCUGGUGGUGGCCAACCCACGCAAGCAGUG 2349

Qy    513 IlePheAlaLysIleIleThrMetPheGly------AsnAsnAspLysThrThrLeuLys 530
Db   2350 CUAUAUGCUAAAGCCAUCAGCACACUCUCGCUCCCAGAAUCACAAUAUAGCAAGCUUCGA 2409

Qy    531 HisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSer 550
Db   2410 CAUGAUGCGGAUUUCUGGUCAUUGAUGUUCUUCGUGGUUGGUAUCAUUCAGUUUAUCACG 2469

Qy    551 TyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArg 570
Db   2470 CAGUCAACCAAUGGUGCUGCAUUUGCCGUAUGCUCCGAGAGACUUAUUCGUCGCGCGAGA 2529

Qy    571 HisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsn 590
Db   2530 AGCACUGCCUUUCGGACGAUACUCCGUCAAGACAUUGCUUUCUUUGACAAGGAAGAGAAU 2589

Qy    591 SerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThr 610
Db   2590 AGCACCGGCGCUCUGACCUCUUUCCUGUCCACCGAGACGAAGCAUCUCUCCGGUGUUAGC 2649

Qy    611 GlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIle 630
Db   2650 GGUGUGACUCUAGGCACGAUCUUGAUGACCUCCACGACCCUAGGAGCGGCUAUCAUUAUU 2709

Qy    631 SerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAla 650
Db   2710 GCCCUGGCGAUUGGGUGGAAAUUGGCCUUAGUUUGUAUCUCGGUUGUGCCGGUUCUCCUG 2769

Qy    651 ValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGlu 670
Db   2770 GCAUGCGGUUUCUACCGAUUCUAUAUGCUAGCCCAGUUUCAAUCACGCUCCAAGCUUGCU 2829

Qy    671 LeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSer 690
Db   2830 UAUGAGGGAUCUGCAAACUUUGCUUGCGAGGCUACAUCGUCUAUCCGCACAGUUGCGUCA 2889

Qy    691 LeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArg 710
Db   2890 UUAACCCGGGAAAGGGAUGUCUGGGAGAUUUACCAUGCCCAGCUUGACGCACAAGGCAGG 2949

Qy    711 AsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIle 730
Db   2950 ACCAGUCUAAUCUCUGUCUUGAGGUCAUCCCUGUUAUAUGCGUCGUCGCAGGCACUUGUU 3009

Qy    731 TyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMet 750
Db   3010 UUCUUCUGCGUUGCGCUCGGGUUUUGGUACGGAGGGACACUUCUUGGUCACCACGAGUAU 3069

Qy    751 ThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGly 770
Db   3070 GACAUUUUCCGCUUCUUUGUUUGUUUCUCCGAGAUUCUCUUUGGUGCUCAAUCCGCGGGC 3129

Qy    771 LysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPhe 790
Db   3130 ACCGUCUUUUCCUUUGCACCAGACAUGGGCAAGGCGAAGAAUGCGGCCGCCGAAUUCCGA 3189

Qy    791 AlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAsp 810
Db   3190 CGACUGUUCGACCGAAAGCCACAAAUUGAUAACUGGUCUGAAGAGGGCGAGAAGCUCGAA 3249

Qy    811 ThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAsp 830
Db   3250 ACGGUGGAAGGUGAAAUCGAAUUUAGGAACGUGCACUUCAGAUACCCGACCCGCCCAGAA 3309
```

```
Qy     831 ValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheVal 850
                 :::|||||||||||    |||:::::::::    |||:::    ||||||    |||
Db    3310 CAGCCUGUCCUGCGCGGCUUGGACCUGACCGUGAAGCCUGGACAAUAUGUUGCGCUUGUC 3369

Qy     851 GlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspPro 870
           |||   ||||||||||||||||||||::::::    ||||||:::|||    |||||
Db    3370 GGACCCAGCGGUUGUGGCAAGAGUACCACCAUUGCAUUGCUUGAGCGCUUUUACGAUGCG 3429

Qy     871 ValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeu 890
           :::    |||    :::|||   |||||    |||       :::||||||:::
Db    3430 AUUGCCGGGUCCAUCCUUGUUGAUGGGAAGGACAUAAGUAAACUAAAUAUCAACUCCUAC 3489

Qy     891 ArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGlu 910
           ||||||   :::::::::|||   |||||||||   |||:::        :::|||    |||
Db    3490 CGCAGCUUUCUGUCACUGGUCAGCCAGGAGCCGACACUGUACCAGGGCACCAUCAAGGAA 3549

Qy     911 AsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAla 930
           ||||||      |||             ||||||    :::    :::    :::|||
Db    3550 AACAUCUUACUUGGUAUUGUCGAAGAUGACGUACCGGAAGAAUUCUUGAUUAAGGCUUGC 3609

Qy     931 AsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnVal 950
                 |||||||||:::    ||||||       |||||||||    :::||||||    |||
Db    3610 AAGGACGCUAAUAUCUACGACUUCAUCAUGUCGCUCCCGGAGGGCUUUAAUACAGUUGUU 3669

Qy     951 GlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAla 970
           |||   ||||||       ||||||||||||||||||||||||:::||||||||||||||||
Db    3670 GGCAGCAAGGGAGGCAUGUUGUCUGGCGGCCAAAAGCAACGUGUGGCCAUUGCCCGAGCC 3729

Qy     971 LeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAsp 990
           ||||||:::    ||||||||||||||||||||||||||||||||||||||||||:::::::
Db    3730 CUUCUUCGGGAUCCCAAAAUCCUUCUUCUCGAUGAAGCGACGUCAGCCCUCGACUCCGAG 3789

Qy     991 SerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuVal 1010
           ||||||||||||||||||    |||||||||    |||       |||||||||    :::
Db    3790 UCAGAAAAGGUCGUCCAGGCGGCUUUGGAUGCCGCUGCCCGAGGCCGAACCACAAUCGCC 3849

Qy    1011 ValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGly 1030
           |||    ||||||||||||    ||||||    ||||||:::|||    |||             |||
Db    3850 GUUGCACACCGACUCAGCACGAUUCAAAAGGCGGACGUUAUCUAUGUUUUCGACCAAGGC 3909

Qy    1031 LysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLys 1050
           ||||||    |||    |||||||||    ||||||::::::    :::          |||:::::::
Db    3910 AAGAUCGUCGAAAGCGGAACGCACAGCGAACUGGUCCAGAAAAAGGGCCGGUACUACGAG 3969

Qy    1051 LeuValAsnAlaGlnSerVal 1057
           |||||||||    ||||||:::
Db    3970 CUGGUCAACUUGCAGAGCUUG 3990
```

RESULT 15
US-09-328-320-1
; Sequence 1, Application US/09328320
; Patent No. 6228615
; GENERAL INFORMATION:
;   APPLICANT: Skatrud, Paul L.
;   APPLICANT: de Waard, Maarten A.
;   APPLICANT: Peery, Robert B.
;   APPLICANT: Andrade, Alan C.
;   TITLE OF INVENTION: Multiple Drug Resistance Gene atrD of
;   TITLE OF INVENTION: Aspergillus Nidulans
;   NUMBER OF SEQUENCES: 3
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Eli Lilly and Company
;   STREET: Lilly Corporate Center
;   CITY: Indianapolis
;   STATE: Indiana
;   COUNTRY: U.S.
;   ZIP: 46285
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/328,320
;   FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/996,545
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Webster, Thomas D.
;   REGISTRATION NUMBER: 39,872
;   REFERENCE/DOCKET NUMBER: X-11766
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 317-276-3334
;   TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4002 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..4002
US-09-328-320-1

```
Alignment Scores:
Pred. No.:               6.76e-222     Length:         4002
Score:                   2067.00       Matches:        447
Percent Similarity:      59.43%        Conservative:   199
Best Local Similarity:   41.12%        Mismatches:     381
Query Match:             38.78%        Indels:         60
DB:                      4             Gaps:           9
```

US-09-873-409-4 (1-1058) x US-09-328-320-1 (1-4002)

```
Qy       2 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 21
           :::|||      :::    |||    |||    :::|||             |||    ||||||||
Db     817 ATCATCAAGTACAGCAAAAAGTCGCTTGACAGCTACGGTGCAGGCGGCACTGTTGCGGAA 876

Qy      22 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 41
           ||||||:::||||||||||||             ||||||    |||:::|||       :::
Db     877 GAGGTCATCAGCTCCATCAGAAATGCCACAGCGTTTGGCACCCAAGACAAGCTTGCGAAG 936

Qy      42 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeu 61

Db     936 ------------------------------------------------------------ 936

Qy      62 SerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 81
                                   :::|||       :::|||    :::|||:::    :::|||
Db     937 --------------------CAGTATGAGGTCCACTTAGACGAAGCTGAGAAATGGGGA 975

Qy      82 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 101
              |||       |||             :::||||||:::::::       |||
Db     976 ACAAAGAACCAGATTGTCATGGGTTTCATGATTGGCGCCATGTTTGGCCTTATGTACTCG 1035

Qy     102 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 121
              |||||||||    ||||||    |||:::       :::::::::|||
Db    1036 AACTACGGTCTTGGCTTCTGGATGGGTTCTCGTTTCCTGGTAGATGGT------GCAGTC 1089

Qy     122 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 141
              :::|||    :::|||    |||       :::::::::       |||:::    :::|||
Db    1090 GATGTGGGTGATATTCTCACAGTTCTCATGGCCATCTTGATCGGATCGTTCTCCTTGGGG 1149

Qy     142 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 161
                       |||:::    :::    |||       |||       ||||||       ||||||
Db    1150 AACGTTAGTCCAAATGCTCAAGCATTTACAAACGCTGTGGCCGCGGCCGCAAAGATATTT 1209

Qy     162 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 181
              ||||||::::::       :::|||       |||       |||             :::
Db    1210 GGAACGATCGATCGCCAGTCCCCATTAGATCCATATTCGAACGAAGGGAAGACGCTCGAC 1269
```

```
Qy    182 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 201
Db   1270 CATTTTGAGGGCCACATTGAGTTACGCAATGTCAAGCATATTTACCCATCTAGACCCGAG 1329

Qy    202 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 221
Db   1330 GTCACCGTCATGGAGGATGTTTCTCTGTCAATGCCCGCTGGAAAAACAACCGCTTTAGTC 1389

Qy    222 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 241
Db   1390 GGCCCCTCTGGCTCTGGAAAAAGTACGGTGGTCGGCTTGGTTGAGCGATTCTACATGCCT 1449

Qy    242 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 261
Db   1450 GTTCGCGGTACGGTTTTGCTGGATGGCCATGACATCAAGGACCTCAATCTCCGCTGGCTT 1509

Qy    262 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 281
Db   1510 CGCCAACAGATCTCTTTGGTTAGCCAGGAGCCTGTTCTTTTTGGCACGACGATTTATAAG 1569

Qy    282 AsnIleLysTyrGly---------------------------ArgAspAspValThrAsp 292
Db   1570 AATATTAGGCACGGTCTCATCGGCACAAAGTACGAGAATGAATCCGAGGATAAGGTCCGG 1629

Qy    293 GluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhePro 312
Db   1630 GAACTCATCGAGAACGCGGCAAAAATGGCGAATGCTCATGACTTTATTACTGCCTTGCCT 1689

Qy    313 AsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGln 332
Db   1690 GAAGGTTATGAGACCAATGTTGGGCAGCGTGGCTTTCTCCTTTCAGGTGGCCAGAAACAG 1749

Qy    333 ArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAla 352
Db   1750 CGCATTGCAATCGCCCGTGCCGTTGTTAGTGACCCAAAAATCCTGCTCCTGGATGAAGCT 1809

Qy    353 ThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSer 372
Db   1810 ACTTCGGCCTTGGACACAAAATCCGAAGGCGTGGTTCAAGCAGCTTTGGAGAGGGCAGCT 1869

Qy    373 LysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeu 392
Db   1870 GAAGGCCGAACTACTATTGTGATCGCTCATCGCCTTTCCACGATCAAAACGGCGCACAAC 1929

Qy    393 IleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAla 412
Db   1930 ATTGTGGTTCTGGTCAATGGCAAAATTGCTGAACAAGGAACTCACGATGAATTGGTTGAC 1989

Qy    413 LysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGln 432
Db   1990 CGCGGAGGCGCTTATCGCAAACTTGTGGAGGCTCAACGTATCAATGAACAGAAGGAAGCT 2049

Qy    433 MetGluSerMetThrTyrSerThrGluArgLysThrAsnSer------LeuProLeuHis 450
Db   2050 GACGCCTTGGAGGACGCCGACGCTGAGGATCTCACGAATGCAGATATTGCCAAAATCAAA 2109

Qy    451 SerValLysSerIleLysSerAspPhe------------------IleAspLys------ 462
Db   2110 ACTGCGTCAAGCGCATCATCCGATCTCGACGGAAAACCCACAACCATTGACCGCACGGGC 2169

Qy    463 AlaGluGluSerThrGlnSerLysGluIleSer------------------LeuProGlu 476
Db   2170 ACCCACAAGTCTGTTTCCAGCGCGATTCTTTCTAAAAGACCCCCCGAAACAACTCCGAAA 2229

Qy    477 ValSerLeuLeuLysIleLeuLys------------LeuAsnLysProGluTrpProPhe 492
Db   2230 TACTCATTATGGACGCTGCTCAAATTTGTTGCTTCCTTCAACCGCCCTGAAATCCCGTAC 2289

Qy    493 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 512
Db   2290 ATGCTCATCGGTCTTGTCTTCTCAGTGTTAGCTGGTGGTGGCCAACCCACGCAAGCAGTG 2349

Qy    513 IlePheAlaLysIleIleThrMetPheGly------AsnAsnAspLysThrThrLeuLys 530
Db   2350 CTATATGCTAAAGCCATCAGCACACTCTCGCTCCCAGAATCACAATATAGCAAGCTTCGA 2409

Qy    531 HisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSer 550
Db   2410 CATGATGCGGATTTCTGGTCATTGATGTTCTTCGTGGTTGGTATCATTCAGTTTATCACG 2469

Qy    551 TyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArg 570
Db   2470 CAGTCAACCAATGGTGCTGCATTTGCCGTATGCTCCGAGAGACTTATTCGTCGCGCGAGA 2529

Qy    571 HisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsn 590
Db   2530 AGCACTGCCTTTCGGACGATACTCCGTCAAGACATTGCTTTCTTTGACAAGGAAGAGAAT 2589

Qy    591 SerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThr 610
Db   2590 AGCACCGGCGCTCTGACCTCTTTCCTGTCCACCGAGACGAAGCATCTCTCCGGTGTTAGC 2649

Qy    611 GlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIle 630
Db   2650 GGTGTGACTCTAGGCACGATCTTGATGACCTCCACGACCCTAGGAGCGGCTATCATTATT 2709

Qy    631 SerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAla 650
Db   2710 GCCCTGGCGATTGGGTGGAAATTGGCCTTAGTTTGTATCTCGGTTGTGCCGGTTCTCCTG 2769

Qy    651 ValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGlu 670
Db   2770 GCATGCGGTTTCTACCGATTCTATATGCTAGCCCAGTTTCAATCACGCTCCAAGCTTGCT 2829

Qy    671 LeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSer 690
Db   2830 TATGAGGGATCTGCAAACTTTGCTTGCGAGGCTACATCGTCTATCCGCACAGTTGCGTCA 2889

Qy    691 LeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArg 710
Db   2890 TTAACCCGGGAAAGGGATGTCTGGGAGATTTACCATGCCCAGCTTGACGCACAAGGCAGG 2949

Qy    711 AsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIle 730
Db   2950 ACCAGTCTAATCTCTGTCTTGAGGTCATCCCTGTTATATGCGTCGTCGCAGGCACTTGTT 3009

Qy    731 TyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMet 750
Db   3010 TTCTTCTGCGTTGCGCTCGGGTTTTGGTACGGAGGGACACTTCTTGGTCACCACGAGTAT 3069

Qy    751 ThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGly 770
Db   3070 GACATTTTCCGCTTCTTTGTTTGTTTCTCCGAGATTCTCTTTGGTGCTCAATCCGCGGGC 3129

Qy    771 LysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPhe 790
Db   3130 ACCGTCTTTTCCTTTGCACCAGACATGGGCAAGGCGAAGAATGCGGCCGCCGAATTCCGA 3189

Qy    791 AlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAsp 810
Db   3190 CGACTGTTCGACCGAAAGCCACAAATTGATAACTGGTCTGAAGAGGGCGAGAAGCTCGAA 3249

Qy    811 ThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAsp 830
Db   3250 ACGGTGGAAGGTGAAATCGAATTTAGGAACGTGCACTTCAGATACCCGACCCGCCCAGAA 3309

Qy    831 ValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheVal 850
Db   3310 CAGCCTGTCCTGCGCGGCTTGGACCTGACCGTGAAGCCTGGACAATATGTTGCGCTTGTC 3369

Qy    851 GlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspPro 870
Db   3370 GGACCCAGCGGTTGTGGCAAGAGTACCACCATTGCATTGCTTGAGCGCTTTTACGATGCG 3429

Qy    871 ValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeu 890
```

```
Db    3430 ATTGCCGGGTCCATCCTTGTTGATGGGAAGGACATAAGTAAACTAAATATCAACTCCTAC 3489

Qy     891 ArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGlu 910
           ||||||   :::::::::|||   |||||||||   |||:::      :::|||   |||
Db    3490 CGCAGCTTTCTGTCACTGGTCAGCCAGGAGCCGACACTGTACCAGGGCACCATCAAGGAA 3549

Qy     911 AsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAla 930
           ||||||      |||            ||||||   :::   :::   :::|||
Db    3550 AACATCTTACTTGGTATTGTCGAAGATGACGTACCGGAAGAATTCTTGATTAAGGCTTGC 3609

Qy     931 AsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnVal 950
                 |||||||||:::   ||||||      |||||||||:::||||||      |||
Db    3610 AAGGACGCTAATATCTACGACTTCATCATGTCGCTCCCGGAGGGCTTTAATACAGTTGTT 3669

Qy     951 GlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAla 970
           |||   ||||||      ||||||||||||||||||||||||:::|||||||||||||||
Db    3670 GGCAGCAAGGGAGGCATGTTGTCTGGCGGCCAAAAGCAACGTGTGGCCATTGCCCGAGCC 3729

Qy     971 LeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAsp 990
           ||||||:::   ||||||||||||||||||||||||||||||||||||||||||::::::
Db    3730 CTTCTTCGGGATCCCAAAATCCTTCTTCTCGATGAAGCGACGTCAGCCCTCGACTCCGAG 3789

Qy     991 SerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuVal 1010
           ||||||||||||||||||   |||||||||   |||      |||||||||   :::
Db    3790 TCAGAAAAGGTCGTCCAGGCGGCTTTGGATGCCGCTGCCCGAGGCCGAACCACAATCGCC 3849

Qy    1011 ValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGly 1030
           |||   ||||||||||||   ||||||   ||||||:::|||   |||         |||
Db    3850 GTTGCACACCGACTCAGCACGATTCAAAAGGCGGACGTTATCTATGTTTTCGACCAAGGC 3909

Qy    1031 LysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLys 1050
           ||||||   |||   |||||||||   ||||||::::::   :::      |||::::::
Db    3910 AAGATCGTCGAAAGCGGAACGCACAGCGAACTGGTCCAGAAAAAGGGCCGGTACTACGAG 3969

Qy    1051 LeuValAsnAlaGlnSerVal 1057
           |||||||||   ||||||:::
Db    3970 CTGGTCAACTTGCAGAGCTTG 3990

Search completed: March 31, 2003, 15:27:41
Job time : 409.939 secs
```

```
                    GenCore version 5.1.4_p5_4578
              Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 31, 2003, 13:52:37 ; Search time 385.104 Seconds
                                          (without alignments)
                                          2337.888 Million cell updates/sec

Title:          US-09-873-409-4
Perfect score:  5330
Sequence:       1 MVISLTSKELSAYSKAGAVA..........QELLRNRDIYFKLVNAQSVQ 1058

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters:      1148742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115424_19290/app_query.fasta_1.7544
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09873409_@CGN_1_1_960_@runat_27032003_115424_19290
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Published_Applications_NA:*
                1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
                2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
                3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
                4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
                5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
                6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
                7:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
                8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
                9:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
                10:  /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
                11:  /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
                12:  /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
                13:  /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
                14:  /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.

                                  SUMMARIES
                 %
Result           Query
  No.    Score   Match Length DB  ID                       Description
---------------------------------------------------------------------------
    1     5328   100.0   3177  10  US-09-873-409-12          Sequence 12, Appl
    2     5328   100.0   3702  10  US-09-873-409-13          Sequence 13, Appl
    3   5159.5    96.8   3621  10  US-09-873-409-14          Sequence 14, Appl
    4   4561.5    85.6   2856  10  US-09-873-409-10          Sequence 10, Appl
    5     3456    64.8   2066  10  US-09-873-409-9           Sequence 9, Appli
    6   3111.5    58.4   4317   9  US-10-044-671-1           Sequence 1, Appli
    7   3106.5    58.3   4369  10  US-09-769-097-1           Sequence 1, Appli
    8   3106.5    58.3   4425  10  US-09-769-097-3           Sequence 3, Appli
    9   3100.5    58.2   3860  10  US-09-866-866A-1          Sequence 1, Appli
   10   3100.5    58.2   3860  10  US-09-866-866A-3          Sequence 3, Appli
   11   3100.5    58.2   8630  10  US-09-306-417-1           Sequence 1, Appli
   12   3100.5    58.2   8630  10  US-09-306-417-2           Sequence 2, Appli
   13   3090.5    58.0   4788  10  US-09-866-866A-7          Sequence 7, Appli
   14     3085    57.9   3924  10  US-09-880-107-2299        Sequence 2299, Ap
   15     3085    57.9   4189  10  US-09-866-866A-5          Sequence 5, Appli
   16     3084    57.9   4643   9  US-10-072-621-2           Sequence 2, Appli
   17   3044.5    57.1   4254  10  US-09-917-800A-1424       Sequence 1424, Ap
   18     3043    57.1   3912  10  US-09-917-800A-1560       Sequence 1560, Ap
   19   2488.5    46.7   5010  10  US-09-917-800A-483        Sequence 483, App
   20     2103    39.5   4175  10  US-09-749-340-3           Sequence 3, Appli
   21   2072.5    38.9   3861   9  US-09-938-842A-2237       Sequence 2237, Ap
   22     2067    38.8   4002   9  US-09-758-828-1           Sequence 1, Appli
   23     2067    38.8   4002   9  US-09-758-828-3           Sequence 3, Appli
   24   2039.5    38.3   3512  10  US-09-749-340-7           Sequence 7, Appli
   25     1951    36.6   3861   9  US-09-938-842A-263        Sequence 263, App
   26     1868    35.0   2021  10  US-09-873-409-15          Sequence 15, Appl
   27   1806.5    33.9   4653   9  US-10-101-388-2           Sequence 2, Appli
   28   1699.5    31.9   1940  10  US-09-873-409-16          Sequence 16, Appl
   29     1617    30.3   2698  10  US-09-749-340-5           Sequence 5, Appli
   30   1538.5    28.9   2681  10  US-09-749-340-8           Sequence 8, Appli
   31   1460.5    27.4   3792   9  US-09-882-694-10          Sequence 10, Appl
   32     1380    25.9   8036   9  US-10-101-388-1           Sequence 1, Appli
   33     1326    24.9   3999   9  US-09-882-694-9           Sequence 9, Appli
   34   1101.5    20.7   1175  10  US-09-873-409-11          Sequence 11, Appl
   35   1032.5    19.4   7296  10  US-09-070-927A-59         Sequence 59, Appl
c  36     1002    18.8   9047  10  US-09-070-927A-12         Sequence 12, Appl
   37    984.5    18.5   3741   9  US-09-738-626-1251        Sequence 1251, Ap
   38    943.5    17.7  12438  10  US-09-070-927A-173        Sequence 173, App
c  39    902.5    16.9   3601  10  US-09-070-927A-255        Sequence 255, App
c  40    901.5    16.9   6415  10  US-09-070-927A-275        Sequence 275, App
   41    868.5    16.3 640681  10  US-09-790-988-1           Sequence 1, Appli
   42    836.5    15.7   1810  10  US-09-749-340-4           Sequence 4, Appli
   43    758.5    14.2   2298   9  US-10-156-239-18          Sequence 18, Appl
   44    758.5    14.2   2298  10  US-09-795-693-18          Sequence 18, Appl
   45    758.5    14.2   3408   9  US-10-156-239-16          Sequence 16, Appl

                                  ALIGNMENTS

RESULT 1
US-09-873-409-12
; Sequence 12, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
;   LENGTH: 3177
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: No. US20020037522A1e
;   LOCATION: (198)..(198)
;   OTHER INFORMATION: n at position 198 represents any nucleotide (A, T, C or G)
US-09-873-409-12

Alignment Scores:
Pred. No.:              0               Length:        3177
Score:                  5328.00         Matches:       1058
Percent Similarity:     100.00%         Conservative:  0
```

```
Best Local Similarity:  100.00%      Mismatches:    0
Query Match:            99.96%       Indels:        0
DB:                     10           Gaps:          0

US-09-873-409-4 (1-1058) x US-09-873-409-12 (1-3177)

Qy      1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCA 60

Qy     21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAGGAGAAAGAACTT 120

Qy     41 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTTGGTTTTATTTCCCCAGTGGCTA 180

Qy     61 LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGAATCTCAAAGATGCAAAGGATTTT 240

Qy     81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAAT 300

Qy    101 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGA 360

Qy    121 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATT 420

Qy    141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATT 480

Qy    161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCT 540

Qy    181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCA 600

Qy    201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTG 660

Qy    221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 720

Qy    241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCAT 780

Qy    261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGT 840

Qy    281 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGG 900

Qy    301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGG 960

Qy    321 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTA 1020

Qy    341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGC 1080

Qy    361 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AAGTCAGCTGTTCAAGCTGCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTA 1140

Qy    381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GCACACCGACTTTCTACTATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATG 1200

Qy    401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTT 1260

Qy    421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACT 1320

Qy    441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATT 1380

Qy    461 AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTA 1440

Qy    481 LysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 AAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCT 1500

Qy    501 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATG 1560

Qy    521 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 TTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTC 1620

Qy    541 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGA 1680

Qy    561 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAG 1740

Qy    581 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCC 1800

Qy    601 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAAT 1860

Qy    621 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTC 1920

Qy    641 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATG 1980

Qy    661 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 ACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAA 2040

Qy    681 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 GCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATG 2100
```

```
Qy    701 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGA 2160

Qy    721 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 AGCTGTTATGCATTCAGCCATGCCTTTATATATTTGCCTATGCAGCAGGGTTTCGATTT 2220

Qy    741 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 GGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACT 2280

Qy    761 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 GCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCC 2340

Qy    781 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 AAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGAC 2400

Qy    801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 AGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAA 2460

Qy    821 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 GTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGT 2520

Qy    841 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 ATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGAAAAAGCACTTCT 2580

Qy    861 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 GTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTG 2640

Qy    881 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 GATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAG 2700

Qy    901 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 CCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTG 2760

Qy    921 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 GTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAA 2820

Qy    941 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 GGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGC 2880

Qy    961 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 CAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTG 2940

Qy    981 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 GATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGAT 3000

Qy   1001 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 AAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAC 3060

Qy   1021 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 GCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAG 3120

Qy   1041 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1058
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 CTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3174
```

RESULT 2

```
US-09-873-409-13
; Sequence 13, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
;   LENGTH: 3702
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: No. US20020037522A1e
;   LOCATION: (723)..(723)
;   OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-13

Alignment Scores:
Pred. No.:              0             Length:        3702
Score:                  5328.00       Matches:       1058
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            99.96%        Indels:        0
DB:                     10            Gaps:          0

US-09-873-409-4 (1-1058) x US-09-873-409-13 (1-3702)

Qy      1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    526 ATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCA 585

Qy     21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    586 GAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAGGAGAAAGAACTT 645

Qy     41 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    646 CAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTTGGTTTTATTTCCCCAGTGGCTA 705

Qy     61 LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    706 CTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGAATCTCAAAGATGCAAAGGATTTT 765

Qy     81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    766 GGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAAT 825

Qy    101 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    826 GGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGA 885

Qy    121 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    886 TATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATT 945

Qy    141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    946 GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATT 1005

Qy    161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1006 TTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCT 1065

Qy    181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1066 GAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCA 1125
```

```
Qy    201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1126 TCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTG 1185

Qy    221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1186 GTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 1245

Qy    241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1246 CCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCAT 1305

Qy    261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1306 TATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGT 1365

Qy    281 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1366 AACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGG 1425

Qy    301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1426 GAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGG 1485

Qy    321 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1486 GAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTA 1545

Qy    341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1546 GTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGC 1605

Qy    361 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1606 AAGTCAGCTGTTCAAGCTGCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTA 1665

Qy    381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1666 GCACACCGACTTTCTACTATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATG 1725

Qy    401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1726 CTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTT 1785

Qy    421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1786 GTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACT 1845

Qy    441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1846 GAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATT 1905

Qy    461 AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1906 GACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTA 1965

Qy    481 LysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1966 AAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCT 2025

Qy    501 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2026 GTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATG 2085

Qy    521 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2086 TTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTC 2145

Qy    541 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2146 GTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGA 2205

Qy    561 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2206 GCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAG 2265

Qy    581 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2266 GATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCC 2325

Qy    601 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2326 ATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAAT 2385

Qy    621 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2386 GCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTC 2445

Qy    641 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2446 CTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATG 2505

Qy    661 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2506 ACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAA 2565

Qy    681 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2566 GCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATG 2625

Qy    701 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2626 TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGA 2685

Qy    721 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2686 AGCTGTTATGCATTCAGCCATGCCTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTT 2745

Qy    741 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2746 GGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACT 2805

Qy    761 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2806 GCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCC 2865

Qy    781 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2866 AAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGAC 2925

Qy    801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2926 AGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAA 2985

Qy    821 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2986 GTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGT 3045

Qy    841 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3046 ATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCT 3105

Qy    861 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3106 GTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTG 3165

Qy    881 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3166 GATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAG 3225

Qy    901 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3226 CCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTG 3285

Qy    921 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    3286 GTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAA 3345

Qy     941 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3346 GGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGC 3405

Qy     961 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3406 CAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTG 3465

Qy     981 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1000
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3466 GATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGAT 3525

Qy    1001 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3526 AAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAC 3585

Qy    1021 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3586 GCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAG 3645

Qy    1041 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1058
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3646 CTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3699
```

RESULT 3
US-09-873-409-14
; Sequence 14, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 3621
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-14

Alignment Scores:

Pred. No.:	0	Length:	3621
Score:	5159.50	Matches:	1031
Percent Similarity:	97.45%	Conservative:	0
Best Local Similarity:	97.45%	Mismatches:	0
Query Match:	96.80%	Indels:	27
DB:	10	Gaps:	1

US-09-873-409-4 (1-1058) x US-09-873-409-14 (1-3621)

```
Qy       1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     526 ATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCA 585

Qy      21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     586 GAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAGGAGAAAGAACTT 645

Qy      41 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60
           |||
Db     646 CAA--------------------------------------------------------- 648

Qy      61 LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80
                                   ||||||||||||||||||||||||||||||||||||
Db     649 ------------------------AGGTATACACAGAATCTCAAAGATGCAAAGGATTTT 684

Qy      81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     685 GGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAAT 744

Qy     101 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     745 GGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGA 804

Qy     121 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     805 TATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATT 864

Qy     141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     865 GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATT 924

Qy     161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     925 TTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCT 984

Qy     181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     985 GAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCA 1044

Qy     201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1045 TCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTG 1104

Qy     221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1105 GTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 1164

Qy     241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1165 CCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCAT 1224

Qy     261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1225 TATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGT 1284

Qy     281 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1285 AACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGG 1344

Qy     301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1345 GAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGG 1404

Qy     321 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1405 GAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTA 1464

Qy     341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1465 GTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGC 1524

Qy     361 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1525 AAGTCAGCTGTTCAAGCTGCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTA 1584

Qy     381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1585 GCACACCGACTTTCTACTATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATG 1644

Qy     401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1645 CTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTT 1704

Qy     421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1705 GTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACT 1764

Qy     441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   1765 GAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATT 1824

Qy    461 AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1825 GACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTA 1884

Qy    481 LysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1885 AAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCT 1944

Qy    501 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1945 GTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATG 2004

Qy    521 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2005 TTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTC 2064

Qy    541 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2065 GTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGA 2124

Qy    561 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2125 GCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAG 2184

Qy    581 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2185 GATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCC 2244

Qy    601 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2245 ATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAAT 2304

Qy    621 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2305 GCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTC 2364

Qy    641 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2365 CTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATG 2424

Qy    661 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2425 ACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAA 2484

Qy    681 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2485 GCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATG 2544

Qy    701 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2545 TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGA 2604

Qy    721 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2605 AGCTGTTATGCATTCAGCCATGCCTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTT 2664

Qy    741 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2665 GGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACT 2724

Qy    761 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2725 GCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCC 2784

Qy    781 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2785 AAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGAC 2844

Qy    801 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2845 AGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAA 2904

Qy    821 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2905 GTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGT 2964

Qy    841 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2965 ATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCT 3024

Qy    861 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3025 GTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTG 3084

Qy    881 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3085 GATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAG 3144

Qy    901 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3145 CCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTG 3204

Qy    921 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3205 GTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAA 3264

Qy    941 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3265 GGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGC 3324

Qy    961 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3325 CAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTG 3384

Qy    981 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3385 GATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGAT 3444

Qy   1001 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3445 AAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAC 3504

Qy   1021 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3505 GCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAG 3564

Qy   1041 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1058
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3565 CTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3618
```

```
RESULT 4
US-09-873-409-10
; Sequence 10, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
;   LENGTH: 2856
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-409-10

Alignment Scores:
Pred. No.:              0              Length:        2856
Score:                  4561.50        Matches:       912
Percent Similarity:     97.96%         Conservative:  0
```

```
Best Local Similarity:  97.96%        Mismatches:    1
Query Match:            85.58%        Indels:       18
DB:                     10            Gaps:          1

US-09-873-409-4 (1-1058) x US-09-873-409-10 (1-2856)

Qy     129 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHis-PheGl 148
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     114 GTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATTGGAGCAGCAGTCCCTCATTATT-- 171

Qy     148 uThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPr 168
                                                            |||||||||||
Db     172 -------------------------------------------------GATAAGAAACC 182

Qy     168 oSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGl 188
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 CAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGA 242

Qy     188 uPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLe 208
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 ATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCT 302

Qy     208 uAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLy 228
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 GAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAA 362

Qy     228 sSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVa 248
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     363 GAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGT 422

Qy     248 lAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVa 268
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     423 GGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTGGT 482

Qy     268 lSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAs 288
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     483 TAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGAGA 542

Qy     288 pAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIl 308
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     543 TGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTTAT 602

Qy     308 eMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGl 328
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     603 CATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGG 662

Qy     328 yGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIl 348
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     663 AGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGAT 722

Qy     348 eLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLe 368
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     723 TTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACT 782

Qy     368 uGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleAr 388
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     783 GGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTAGCACACCGACTTTCTACTATTCG 842

Qy     388 gSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAl 408
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     843 AAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATGCTGGCGGAGAAAGGAGCACATGC 902

Qy     408 aGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLy 428
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     903 TGAACTAATGGCAAAACGAGGTCTATATTATTCACTTGTGATGTCACAGGATATTAAAAA 962

Qy     428 sAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPr 448
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     963 AGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCACTTCC 1022

Qy     448 oLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGl 468
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1023 TCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCCACCCA 1082

Qy     468 nSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPr 488
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1083 ATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAACAAGCC 1142

Qy     488 oGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPr 508
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1143 TGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCTGTTCTAAATGGAACTGTTCATCC 1202

Qy     508 oValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrTh 528
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1203 AGTATTTTCCATCATCTTTGCAAAAATTATAACCATGTTTGGAAATAATGATAAAACCAC 1262

Qy     528 rLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPh 548
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1263 ATTAAAGCATGATGCAGAAATTTATTCCATGATATTCGTCATTTTGGGTGTTATTTGCTT 1322

Qy     548 eValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetAr 568
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1323 TGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGAGCAGGGGAAATTTTAACGATGAG 1382

Qy     568 gLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLy 588
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1383 ATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAGGATATTGCCTGGTTTGATGAAAA 1442

Qy     588 sGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGl 608
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1443 GGAAAACAGCACAGGAGGCTTGACAACAATATTAGCCATAGATATAGCACAAATTCAAGG 1502

Qy     608 yAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVa 628
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1503 AGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAATGCAACTAACATGGGACTTTCAGT 1562

Qy     628 lIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVa 648
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1563 TATCATTTCCTTTATATATGGATGGGAGATGACATTCCTGATTCTGAGTATTGCTCCAGT 1622

Qy     648 lLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLy 668
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1623 ACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATGACTGGATTTGCCAACAAAGATAA 1682

Qy     668 sGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIl 688
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1683 GCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAGAATATACGTACTAT 1742

Qy     688 eValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGl 708
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1743 AGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATGTATGAAGAGATGCTTCAGACTCA 1802

Qy     708 nHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAl 728
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1803 ACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTATGCATTCAGCCATGC 1862

Qy     728 aPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGl 748
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1863 CTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTTGGAGCCTATTTAATTCAAGCTGG 1922

Qy     748 yArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAl 768
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1923 ACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACTGCAATTGCATATGGAGCTATGGC 1982

Qy     768 aIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHi 788
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1983 CATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCCAAAGCCAAATCGGGGGCTGCGCA 2042

Qy     788 sLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLy 808
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2043 TCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGACAGCCGCAGTCAAGAAGGGAAAAA 2102

Qy     808 sProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysAr 828
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2103 GCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTCTTCTATCCATGTCG 2162
```

```
Qy    828 gProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAl 848
Db   2163 CCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGC 2222

Qy    848 aPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTy 868
Db   2223 ATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTA 2282

Qy    868 rAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGl 888
Db   2283 TGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTGGATGCAAAAGAATTGAATGTACA 2342

Qy    888 nTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIl 908
Db   2343 GTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAGCCTGTGCTCTTCAACTGCAGCAT 2402

Qy    908 eAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGl 928
Db   2403 TGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTGGTGCCATTAGATGAGATCAAAGA 2462

Qy    928 uAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnTh 948
Db   2463 AGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAAGGTCTCCCTGAGAAATACAACAC 2522

Qy    948 rGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAl 968
Db   2523 ACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCTATTGC 2582

Qy    968 aArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAs 988
Db   2583 AAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGA 2642

Qy    988 pAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCy 1008
Db   2643 TAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGATAAAGCCAGGACGGGAAGGACATG 2702

Qy   1008 sLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHi 1028
Db   2703 CCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGGTTCTGCA 2762

Qy   1028 sAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTy 1048
Db   2763 CAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAGCTCCTGAGAAATCGAGACATATA 2822

Qy   1048 rPheLysLeuValAsnAlaGlnSerValGln 1058
Db   2823 TTTTAAGTTAGTGAATGCACAGTCAGTGCAG 2853

RESULT 5
US-09-873-409-9
; Sequence 9, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 9
;   LENGTH: 2066
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-409-9

Alignment Scores:
Pred. No.:              3.38e-314      Length:         2066
Score:                  3456.00        Matches:        687
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0

Query Match:            64.84%         Indels:         0
DB:                     10             Gaps:           0

US-09-873-409-4 (1-1058) x US-09-873-409-9 (1-2066)

Qy    372 SerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAsp 391
Db      3 AGCAAAGGTCGGACTACAATCGTGGTAGCACACCGACTTTCTACTATTCGAAGTGCAGAT 62

Qy    392 LeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMet 411
Db     63 TTGATTGTGACCCTAAAGGATGGAATGCTGGCGGAGAAAGGAGCACATGCTGAACTAATG 122

Qy    412 AlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGlu 431
Db    123 GCAAAACGAGGTCTATATTATTCACTTGTGATGTCACAGGATATTAAAAAAGCTGATGAA 182

Qy    432 GlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSer 451
Db    183 CAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCACTTCCTCTGCACTCT 242

Qy    452 ValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGlu 471
Db    243 GTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCCACCCAATCTAAAGAG 302

Qy    472 IleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpPro 491
Db    303 ATAAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAACAAGCCTGAATGGCCT 362

Qy    492 PheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSer 511
Db    363 TTTGTGGTTCTGGGGACATTGGCTTCTGTTCTAAATGGAACTGTTCATCCAGTATTTTCC 422

Qy    512 IleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHis 531
Db    423 ATCATCTTTGCAAAAATTATAACCATGTTTGGAAATAATGATAAAACCACATTAAAGCAT 482

Qy    532 AspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyr 551
Db    483 GATGCAGAAATTTATTCCATGATATTCGTCATTTTGGGTGTTATTTGCTTTGTCAGTTAT 542

Qy    552 PheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHis 571
Db    543 TTCATGCAGGGATTATTTTACGGCAGAGCAGGGGAAATTTTAACGATGAGATTAAGACAC 602

Qy    572 LeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSer 591
Db    603 TTGGCCTTCAAAGCCATGTTATATCAGGATATTGCCTGGTTTGATGAAAAGGAAAACAGC 662

Qy    592 ThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGly 611
Db    663 ACAGGAGGCTTGACAACAATATTAGCCATAGATATAGCACAAATTCAAGGAGCAACAGGT 722

Qy    612 SerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSer 631
Db    723 TCCAGGATTGGCGTCTTAACACAAAATGCAACTAACATGGGACTTTCAGTTATCATTTCC 782

Qy    632 PheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaVal 651
Db    783 TTTATATATGGATGGGAGATGACATTCCTGATTCTGAGTATTGCTCCAGTACTTGCCGTG 842

Qy    652 ThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeu 671
Db    843 ACAGGAATGATTGAAACCGCAGCAATGACTGGATTTGCCAACAAAGATAAGCAAGAACTT 902

Qy    672 LysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeu 691
Db    903 AAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAGAATATACGTACTATAGTGTCATTA 962

Qy    692 ThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsn 711
Db    963 ACAAGGAAAAAAGCCTTCGAGCAAATGTATGAAGAGATGCTTCAGACTCAACACAGAAAT 1022
```

```
Qy      712 ThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyr 731
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1023 ACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTATGCATTCAGCCATGCCTTTATATAT 1082

Qy      732 PheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThr 751
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1083 TTTGCCTATGCAGCAGGGTTTCGATTTGGAGCCTATTTAATTCAAGCTGGACGAATGACC 1142

Qy      752 ProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLys 771
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1143 CCAGAGGGCATGTTCATAGTTTTTACTGCAATTGCATATGGAGCTATGGCCATCGGAAAA 1202

Qy      772 ThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAla 791
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1203 ACGCTCGTTTTGGCTCCTGAATATTCCAAAGCCAAATCGGGGGCTGCGCATCTGTTTGCC 1262

Qy      792 LeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThr 811
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1263 TTGTTGGAAAAGAAACCAAATATAGACAGCCGCAGTCAAGAAGGGAAAAAGCCAGACACA 1322

Qy      812 CysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspVal 831
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1323 TGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTCTTCTATCCATGTCGCCCAGATGTT 1382

Qy      832 PheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGly 851
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1383 TTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGCATTTGTGGGG 1442

Qy      852 SerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProVal 871
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1443 AGCAGCGGCTGTGGGAAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTG 1502

Qy      872 GlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArg 891
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1503 CAAGGACAAGTGCTGTTTGATGGTGTGGATGCAAAAGAATTGAATGTACAGTGGCTCCGT 1562

Qy      892 SerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsn 911
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1563 TCCCAAATAGCAATCGTTCCTCAAGAGCCTGTGCTCTTCAACTGCAGCATTGCTGAGAAC 1622

Qy      912 IleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsn 931
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1623 ATCGCCTATGGTGACAACAGCCGTGTGGTGCCATTAGATGAGATCAAAGAAGCCGCAAAT 1682

Qy      932 AlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGly 951
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1683 GCAGCAAATATCCATTCTTTTATTGAAGGTCTCCCTGAGAAATACAACACACAAGTTGGA 1742

Qy      952 LeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeu 971
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1743 CTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTT 1802

Qy      972 LeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSer 991
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1803 CTCCAAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGATAATGACAGT 1862

Qy      992 GluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValVal 1011
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1863 GAGAAGGTGGTTCAGCATGCCCTTGATAAAGCCAGGACGGGAAGGACATGCCTAGTGGTC 1922

Qy     1012 ThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLys 1031
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1923 ACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGGTTCTGCACAATGGAAAG 1982

Qy     1032 IleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeu 1051
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1983 ATAAAGGAACAAGGAACTCATCAAGAGCTCCTGAGAAATCGAGACATATATTTTAAGTTA 2042

Qy     1052 ValAsnAlaGlnSerValGln 1058
            |||||||||||||||||||||
Db     2043 GTGAATGCACAGTCAGTGCAG 2063
```

RESULT 6

```
US-10-044-671-1
; Sequence 1, Application US/10044671
; Patent No. US20020177147A1
; GENERAL INFORMATION:
;  APPLICANT: Washington State University Research Foundation
;  APPLICANT:  Mealey, Katrina
;  APPLICANT:  Bentjen, Steven
;  TITLE OF INVENTION: MDR1 VARIANTS AND METHODS FOR THEIR USE
;  FILE REFERENCE: 4630-61733
;  CURRENT APPLICATION NUMBER: US/10/044,671
;  CURRENT FILING DATE:  2002-01-10
;  PRIOR APPLICATION NUMBER: US 60/261,578
;  PRIOR FILING DATE: 2001-01-12
;  PRIOR APPLICATION NUMBER: US 60/314,829
;  PRIOR FILING DATE: 2001-08-24
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 4317
;   TYPE: DNA
;   ORGANISM: Canis familiaris
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (70)..(3912)
;   OTHER INFORMATION:
US-10-044-671-1

Alignment Scores:
Pred. No.:              1.9e-281       Length:        4317
Score:                  3111.50        Matches:       600
Percent Similarity:     75.02%         Conservative:  202
Best Local Similarity:  56.13%         Mismatches:    220
Query Match:            58.38%         Indels:        47
DB:                     9              Gaps:          6

US-09-873-409-4 (1-1058) x US-10-044-671-1 (1-4317)

Qy        4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
              |||   |||   ||||||||   ||||||:::||||||||||||||||||||||||||
Db      784 TCATTTACTGATAAAGAACTCTTGGCCTATGCAAAAGCTGGAGCAGTAGCTGAAGAAGTC 843

Qy       24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
            |||::::::||||||||||||||||||||||        |||:::|||||||||:::
Db      844 TTAGCAGCAATCAGAACTGTGATTGCCTTTGGAGGACAAAAGAAAGAACTTGAA------ 897

Qy       44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63

Db      897 ------------------------------------------------------------ 897

Qy       64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
                           |||||    :::||||||:::::|||||         ||||||||
Db      898 ---------------AGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGATAAAG 942

Qy       84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
            :::    |||    :::    :::|||:::||||||    :::       :::       :::|||
Db      943 AAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCATCATAT 1002

Qy      104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
               ||||||||||||||||||||||||||:::|||:::    |||      |||||||||
Db     1003 GCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TATACTATT 1056

Qy      124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
            |||   ||||||   ||||||||||||||:::         ::::::   ||||||   |||
Db     1057 GGACAGGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGACAGGCA 1116

Qy      144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
               |||      |||   ||||||   |||||||||||||||:::    ||||||::::::
Db     1117 TCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1176

Qy      164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
            ||||||   |||||||||||||||:::::::|||   :::|||:::||||||::::::|||
Db     1177 ATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGATAATATT 1236
```

```
Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db   1237 AAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAAGTTAAG 1296

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db   1297 ATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTTGGGAAC 1356

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db   1357 AGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCCACAGAT 1416

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db   1417 GGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTTCGGGAA 1476

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db   1477 ATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAAAACATT 1536

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db   1537 CGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAAGCCAAT 1596

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db   1597 GCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAGAGAGGG 1656

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db   1657 GCCCGGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTTCGCAAC 1716

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db   1717 CCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAAGCAGTG 1776

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db   1777 GTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCTCATCGT 1836

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db   1837 TTGTCTACAGTTCGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATTGTGGAG 1896

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   1897 AAAGGAAATCATGATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTCACAATG 1956

Qy    424 GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 443
Db   1957 CAG---ACAAGAGGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAAAGTGAA 2013

Qy    444 ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--------- 460
Db   2014 AGTGATGCCTTGGAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAAAGAAGA 2070

Qy    461 -----------------------------------AspLysAlaGluGluSerThrGln 468
Db   2071 TCAACTCGCAGGAGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACAAAAGAG 2130

Qy    469 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 488
Db   2131 GACTTGAATGAGAATGTTCCTTCAGTTTCCTTCTGGAGGATTCTGAAGCTGAACTCAACT 2190

Qy    489 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 508
Db   2191 GAATGGCCTTATTTTGTGGTTGGTATATTTTGTGCTATTATAAACGGAGGCCTGCAACCA 2250

Qy    509 ValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLysThr 527
Db   2251 GCATTTTCAATAATATTTTCAAGGATTATAGGGATCTTTACCCGAGATGAGGATCCTGAA 2310

Qy    528 ThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 547
Db   2311 ACAAAACGACAGAATAGTAACATGTTTTCTGTATTGTTTCTAGTCCTTGGAATTATTTCT 2370

Qy    548 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 567
Db   2371 TTTATTACATTTTTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTCACTAAG 2430

Qy    568 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 587
Db   2431 CGGCTTCGATACATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTTGATGAC 2490

Qy    588 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 607
Db   2491 CCTAAAAACACCACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAAGTTAAA 2550

Qy    608 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 627
Db   2551 GGGGCTATAGGTTCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGGACAGGC 2610

Qy    628 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 647
Db   2611 ATTATTATATCCTTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATTGTACCC 2670

Qy    648 ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp 667
Db   2671 ATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGAT 2730

Qy    668 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 687
Db   2731 AAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTCCGAACT 2790

Qy    688 IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 707
Db   2791 GTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTGCAAGTA 2850

Qy    708 GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis 727
Db   2851 CCATACAGAAACTCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATCACCCAG 2910

Qy    728 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla 747
Db   2911 GCAATGATGTATTTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTGGCAAAT 2970

Qy    748 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 767
Db   2971 GAGTTCATGAACTTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGTGCCATG 3030

Qy    768 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 787
Db   3031 GCAGTGGGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCAGCAGCC 3090

Qy    788 HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys 807
Db   3091 CACGTCATCATGATCATTGAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCACGGCCTC 3150

Qy    808 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys 827
Db   3151 AAGCCAAATACGTTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTATCCCACT 3210

Qy    828 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 847
Db   3211 CGACCAGACATCCCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAGACGCTG 3270

Qy    848 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 867
Db   3271 GCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAGCGCTTC 3330

Qy    868 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 887
Db   3331 TATGACCCCTTGGCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTGAATGTC 3390

Qy    888 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 907
```

```
Db    3391 CAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGACTGCAGC 3450

Qy     908 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 927

Db    3451 ATTGCCGAGAACATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAGATTATG 3510

Qy     928 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn 947

Db    3511 CAGGCAGCCAAGGAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAATACAAC 3570

Qy     948 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 967

Db    3571 ACCAGAGTAGGAGACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATTGCCATA 3630

Qy     968 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 987

Db    3631 GCTCGCGCTCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCTACATCAGCTCTG 3690

Qy     988 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1007

Db    3691 GATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACC 3750

Qy    1008 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1027

Db    3751 TGCATTGTGATCGCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTGGTGTTT 3810

Qy    1028 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1047

Db    3811 CAGAATGGCAAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGGCTCAGAAAGGCATC 3870

Qy    1048 TyrPheLysLeuValAsnAlaGlnSer 1056

Db    3871 TATTTTTCCATGATCAGTGTCCAGGCT 3897

RESULT 7
US-09-769-097-1
; Sequence 1, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
;  APPLICANT: Kimberly Anne Brun
;  APPLICANT:  Richard James Chenery
;  APPLICANT:  Harma Ellens
;  APPLICANT:  John Anthony Feild
;  APPLICANT:  Lin Yue
;  TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
;  TITLE OF INVENTION:  ENCODING RAT MDR1A AND SCREENING METHODS THEREOF
;  FILE REFERENCE: GP-50009-C2
;  CURRENT APPLICATION NUMBER: US/09/769,097
;  CURRENT FILING DATE:  2001-01-25
;  PRIOR APPLICATION NUMBER: 09/208,809
;  PRIOR FILING DATE: 1998-12-09
;  PRIOR APPLICATION NUMBER: 09/156,800
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: US99/20770
;  PRIOR FILING DATE: 1999-09-10
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 4369
;   TYPE: DNA
;   ORGANISM: RATTUS RATTUS
US-09-769-097-1

Alignment Scores:
Pred. No.:               5.7e-281      Length:          4369
Score:                   3106.50       Matches:         600
Percent Similarity:      75.54%        Conservative:    206
Best Local Similarity:   56.23%        Mismatches:      218
Query Match:             58.28%        Indels:          43
DB:                      10            Gaps:            7

US-09-873-409-4 (1-1058) x US-09-769-097-1 (1-4369)

Qy       4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23

Db    1039 TCATTTACTGATAAGGAACTCCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTC 1098

Qy      24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43

Db    1099 TTAGCAGCCATCAGAACTGTGATTGCCTTTGGAGGACAAAAGAAGGAACTTGAA------ 1152

Qy      44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63

Db    1152 ------------------------------------------------------------ 1152

Qy      64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83

Db    1153 --------------AGGTACAATAACAATTTGGAAGAAGCTAAAAGGCTTGGGATAAAG 1197

Qy      84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103

Db    1198 AAAGCTATCACGGCCAACATTTCCATGGGTGCAGCTTTTCTGCTTATCTATGCATCATAT 1257

Qy     104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123

Db    1258 GCTCTGGCATTCTGGTATGGGACTTCCTTGGTCATCTCAAAAGAA------TACACTATT 1311

Qy     124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143

Db    1312 GGACAAGTGCTCACTGTCTTTTTTTCTGTATTAATTGGAGCATTCAGTGTTGGGCAGGCA 1371

Qy     144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163

Db    1372 TCTCCAAATATTGAAGCCTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTCAGTATA 1431

Qy     164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183

Db    1432 ATTGATAATAAGCCCAGTATAGACAGCTTCTCAAAGAGTGGGCACAAACCCGACAACATA 1491

Qy     184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203

Db    1492 CAAGGAAATTTGGAATTCAAAAATATTCACTTCAGTTACCCGTCTCGAAAAGACGTTCAG 1551

Qy     204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223

Db    1552 ATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTGGTTGGCAAC 1611

Qy     224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243

Db    1612 AGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCATAGAG 1671

Qy     244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263

Db    1672 GGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAA 1731

Qy     264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283

Db    1732 ATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCCACCACAATTGCCGAAAACATT 1791

Qy     284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303

Db    1792 CGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAAT 1851

Qy     304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323

Db    1852 GCCTATGATTTCATCATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAGAGAGGG 1911

Qy     324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343

Db    1912 GCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAAC 1971

Qy     344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363

Db    1972 CCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTG 2031

Qy     364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
```

```
              ||||||||||||||:::||||||    :::||||||||||||||||||:::|||||||||
Db     2032 GTTCAGGCCGCTCTGGATAAGGCTAGAGAAGGCCGGACCACCATTGTGATAGCTCACCGC 2091

Qy      384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
              |||||||||:::|||:::||||||:::|||                   |||::::::   |||
Db     2092 TTGTCTACAGTTCGCAATGCTGACGTCATTGCTGGTTTTGATGGTGGTGTCATTGTGGAG 2151

Qy      404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
              :::|||   |||   |||||||||   ::::::|||:::|||:::   |||||||||:::
Db     2152 CAAGGAAATCATGATGAGCTCATGAGAGAGAAAGGAATTTACTTCAAACTTGTCATGACT 2211

Qy      424 Gln------------AspIleLysLysAlaAspGluGlnMetGluSer------------ 435
              |||             :::|||:::         :::|||      ||||||
Db     2212 CAGACAGCAGGAAATGAAATTGAATTAGGAAATGAAGCTTGTGAATCTAAAGACGGAATT 2271

Qy      436 ------MetThrTyrSerThrGluArgLysThrAsnSerLeu------ProLeuHisSer 451
                    :::         |||:::::::         :::||||||               :::
Db     2272 GATAATGTGGACATGTCTTCAAAAGATTCGGGATCCAGTCTAATAAGAAGAAGATCAACT 2331

Qy      452 ValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLys 470
                 |||||||||:::                         |||   ||||||   ::::::
Db     2332 CGCAAAAGCATCCGTGGGCCACATGATCAAGACGGGGAACTTAGCACCAAAGAGGCTCTG 2391

Qy      471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
              :::         :::|||         |||         :::|||||||||||||||     ||||||
Db     2392 GATGACGACGTACCTCCAGCTTCCTTTTGGCGGATCCTGAAGTTGAATTCAACTGAATGG 2451

Qy      491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
              |||:::      |||:::|||          ::::::::::||||||      :::      |||      |||
Db     2452 CCTTATTTTGTGGTTGGTGTATTTTGTGCCATAATAAATGGAGGCTTGCAACCAGCATTC 2511

Qy      511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeu--- 529
              ||||||||||||:::|||::::::      :::|||            ||||||            :::
Db     2512 TCCATAATATTTCAAAGGTTGTAGGGGTTTTTACAAAAAATGACACCCCTGAAATCCAG 2571

Qy      530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
              :::   ::::::   ::::::|||::::::|||:::|||||||||:::|||      |||:::
Db     2572 CGGCAGAACAGCAACTTGTTTTCTTTATTGTTCCTGATCCTTGGGATCATCTCTTTCATT 2631

Qy      550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
              ::::::|||:::||||||   -   :::|||:::|||||||||||||||||||      ||||||
Db     2632 ACGTTTTTCCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGACTC 2691

Qy      570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
              |||::::::      ||||||:::||||||      |||||||||:::|||||||||:::      :::
Db     2692 CGATACATGGTCTTCAAATCCATGCTGAGACAGGACATAAGCTGGTTTGATGACCCTAAA 2751

Qy      590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
              |||:::||||||   |||||||||   ||||||   |||      ||||||::::::||||||
Db     2752 AACACCACAGGAGCGCTGACCACCAGGCTTGCCAATGACGCTGCTCAAGTGAAAGGGGCT 2811

Qy      610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
              ||||||||||||:::      |||:::|||||||||         |||:::|||         :::|||
Db     2812 ACAGGGTCTAGGCTTGCTGTTATTACCCAGAACATAGCAAATCTTGGGACAGGCATCATC 2871

Qy      630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
              ||||||   ||||||||||||::::::|||   |||:::|||:::|||   |||::::::
Db     2872 ATATCCCTGATCTACGGCTGGCAATTGACACTTTTACTCCTAGCAATTGTTCCCATCATT 2931

Qy      650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
              |||:::      |||::::::|||                  ::::::|||   |||      |||||||||:::
Db     2932 GCTATAGCAGGAGTGGTTGAAATGAAAATGTTGTCTGGACAAGCGCTGAAAGATAAGAAG 2991

Qy      670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
              ||||||:::      :::|||||||||||||||||||||:::||||||      ||||||:::|||
Db     2992 GAACTAGAAGGTTCTGGGAAGATCGCTACAGAAGCAATTGAAAACTTTCGCACTGTCGTC 3051

Qy      690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
              |||||||||||||:::      ||||||   ||||||      :::      ||||||               :::
Db     3052 TCTTTGACTCGGGAGCAGAAGTTTGAAACTATGTATGCCCAGAGCTTGCAGATACCATAC 3111

Qy      710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
              ||||||         |||||||||      :::      |||                  ::::::|||:::      |||
Db     3112 AGAAATGCTTTGAAGAAAGCGCACGTCTTTGGGATCACTTTCTCCTTCACCCAGGCCATG 3171

Qy      730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
              :::|||||||:::|||||||||   |||||||||   |||||||||:::
Db     3172 ATGTATTTCTCCTATGCTGCTTGTTTCCGGTTTGATGCCTACTTGGTGGCACGAGAACTC 3231

Qy      750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
              ||||||      |||      :::      :::||||||:::||||||      :::||||||||||||:::
Db     3232 ATGACATTTGAAAATGTTCTGTTAGTATTCTCAGCTATTGTCTTTGGTGCCATGGCAGTG 3291

Qy      770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
              |||:::                     ||||||:::|||:::|||||||                  |||:::|||:::
Db     3292 GGGCAGGTCAGTTCATTCGCTCCTGACTACGCGAAAGCCAAAGTCTCGGCATCCCACATC 3351

Qy      790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
                    ::::::||||||   |||   |||||||||   |||   ||||||   ||||||
Db     3352 ATCAGGATCATTGAGAAAATCCCTGAGATTGACAGCTACAGCACGGAGGGCTTGAAGCCT 3411

Qy      810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
              :::         |||||||||:::::::|||         |||   |||   ||||||   ||||||
Db     3412 AATATGTTGGAAGGAAATGTGAAATTTAATGGAGTCATGTTCAACTATCCCACCCGACCC 3471

Qy      830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
              ::::::   :::|||:::||||||||||||   :::::::::|||:::|||:::|||
Db     3472 AACATCCCAGTGCTTCAGGGGCTGAGCCTAGAGGTGAAGAAGGGCAGACGCTGGCCCTC 3531

Qy      850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
              |||||||||||||||||||||||||||   ||||||||||||:::|||   ||||||
Db     3532 GTGGGCAGCAGTGGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTTGAGCGCTTCTATGAC 3591

Qy      870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
              |||:::      |||   |||      ||||||      :::   |||:::|||||||||||||||
Db     3592 CCCATGGCCGGAACAGTGTTTCTAGATGGCAAAGAAATAAAGCAACTCAATGTCCAGTGG 3651

Qy      890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
              ||||||:::   :::   ||||||   |||||||||:::||||||:::||||||||||||
Db     3652 CTCCGCGCCCACCTGGGCATTGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATCGCC 3711

Qy      910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
              ||||||||||||||||||||||||||||||||||||      ||||||   :::|||
Db     3712 GAGAACATTGCCTACGGAGACAACAGCCGTGTCGTGTCTCATAAGGAGATCGTGAAGGCA 3771

Qy      930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
              |||      ||||||||||||   ||||||:::   ||||||||||||||||||||:::
Db     3772 GCCAAGGAGGCCAACATCCACCAGTTCATCGACTCACTGCCTGAGAAATACAACACCAGA 3831

Qy      950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
              ||||||   ||||||   |||||||||||||||||||||:::::::::::::::
Db     3832 GTGGGAGACAAAGGGACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATCGCCATCGCGCGC 3891

Qy      970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
              ||||||::::::::::|||   |||||||||||||||||||||||||||||||||
Db     3892 GCCCTCGTCAGACAGCCTCACATCTTACTTCTGGATGAAGCGACATCAGCTCTGGATACG 3951

Qy      990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
              :::||||||||||||||||   ||||||||||||||||   ||||||||||||:::
Db     3952 GAGAGTGAAAAGGTCGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACCTGCATT 4011

Qy     1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
              |||:::   ||||||||||||   ||||||||||||||||||||||||:::   |||
Db     4012 GTGATCGCGCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATTCAGAAC 4071

Qy     1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
              |||::::::||||||   |||||||||:::||||||         :::      |||||||||
Db     4072 GGCCAGGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCCCAGAAAGGCATCTATTTC 4131

Qy     1050 LysLeuValAsnAlaGlnSer 1056
                    :::|||:::      |||:::
Db     4132 TCGATGGTCAGTGTGCAGGCT 4152
```

RESULT 8
US-09-769-097-3
; Sequence 3, Application US/09769097

```
; Patent No. US20020055128A1
; GENERAL INFORMATION:
;  APPLICANT: Kimberly Anne Brun
;  APPLICANT:  Richard James Chenery
;  APPLICANT:  Harma Ellens
;  APPLICANT:  John Anthony Feild
;  APPLICANT:  Lin Yue
;  TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
;  TITLE OF INVENTION:  ENCODING RAT MDR1A AND SCREENING METHODS THEREOF
;  FILE REFERENCE: GP-50009-C2
;  CURRENT APPLICATION NUMBER: US/09/769,097
;  CURRENT FILING DATE:  2001-01-25
;  PRIOR APPLICATION NUMBER: 09/208,809
;  PRIOR FILING DATE: 1998-12-09
;  PRIOR APPLICATION NUMBER: 09/156,800
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: US99/20770
;  PRIOR FILING DATE: 1999-09-10
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 4425
;   TYPE: DNA
;   ORGANISM: RATTUS RATTUS
US-09-769-097-3

Alignment Scores:
Pred. No.:              5.81e-281     Length:        4425
Score:                  3106.50       Matches:       600
Percent Similarity:     75.54%        Conservative:  206
Best Local Similarity:  56.23%        Mismatches:    218
Query Match:            58.28%        Indels:        43
DB:                     10            Gaps:          7

US-09-873-409-4 (1-1058) x US-09-769-097-3 (1-4425)

Qy      4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
Db   1039 TCATTTACTGATAAGGAACTCCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTC 1098

Qy     24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db   1099 TTAGCAGCCATCAGAACTGTGATTGCCTTTGGAGGACAAAAGAAGGAACTTGAA------ 1152

Qy     44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db   1152 ------------------------------------------------------------ 1152

Qy     64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db   1153 ---------------AGGTACAATAACAATTTGGAAGAAGCTAAAAGGCTTGGGATAAAG 1197

Qy     84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db   1198 AAAGCTATCACGGCCAACATTTCCATGGGTGCAGCTTTTCTGCTTATCTATGCATCATAT 1257

Qy    104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db   1258 GCTCTGGCATTCTGGTATGGGACTTCCTTGGTCATCTCAAAAGAA------TACACTATT 1311

Qy    124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db   1312 GGACAAGTGCTCACTGTCTTTTTTTCTGTATTAATTGGAGCATTCAGTGTTGGCAGGCA 1371

Qy    144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db   1372 TCTCCAAATATTGAAGCCTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTCAGTATA 1431

Qy    164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db   1432 ATTGATAATAAGCCCAGTATAGACAGCTTCTCAAAGAGTGGGCACAAACCCGACAACATA 1491

Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db   1492 CAAGGAAATTTGGAATTCAAAAATATTCACTTCAGTTACCCGTCTCGAAAAGACGTTCAG 1551

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db   1552 ATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTGGTTGGCAAC 1611

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db   1612 AGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCATAGAG 1671

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db   1672 GGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAA 1731

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db   1732 ATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCCACCACAATTGCCGAAAACATT 1791

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db   1792 CGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAAT 1851

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db   1852 GCCTATGATTTCATCATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAGAGAGGG 1911

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db   1912 GCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAAC 1971

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db   1972 CCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTG 2031

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db   2032 GTTCAGGCCGCTCTGGATAAGGCTAGAGAAGGCCGGACCACCATTGTGATAGCTCACCGC 2091

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db   2092 TTGTCTACAGTTCGCAATGCTGAYGTCATTGCTGGTTTTGATGGTGGTGTCATTGTGGAG 2151

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   2152 CAAGGAAATCATGATGAGCTCATGAGAGAGAAAGGAATTTACTTCAAACTTGTCATGACT 2211

Qy    424 Gln------------AspIleLysLysAlaAspGluGlnMetGluSer------------ 435
Db   2212 CAGACAGCAGGAAATGAAATTGAATTAGGAAATGAAGCTTGTGAATCTAAAGAYGGAATT 2271

Qy    436 ------MetThrTyrSerThrGluArgLysThrAsnSerLeu------ProLeuHisSer 451
Db   2272 GATAATGTGGACATGTCTTCAAAAGATTCRGGATCCAGTCTAATAAGAAGAAGATCAACT 2331

Qy    452 ValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLys 470
Db   2332 CGCAAAAGCATCCGTGGGCCACATGATCAAGACGGGGAACTTAGCACCAAAGAGGCTCTG 2391

Qy    471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
Db   2392 GATGACGACGTACCTCCAGCTTCCTTTGGCGGATCCTGAAGTTGAATTCAACTGAATGG 2451

Qy    491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db   2452 CCTTATTTTGTGGTTGGTGTATTTTGTGCCATAATAAATGGAGGCTTGCAACCAGCATTC 2511

Qy    511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeu--- 529
Db   2512 TCCATAATATTTTCAAAGGTTGTAGGGGTTTTTACAAAAAATGACACCCCTGAAATCCAG 2571

Qy    530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
```

```
Db   2572 CGGCAGAACAGCAACTTGTTTTCTTTATTGTTCCTGATCCTTGGGATCATCTCTTTCATT 2631
Qy    550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db   2632 ACGTTTTTCCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGACTC 2691
Qy    570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
Db   2692 CGATACATGGTCTTCAAATCCATGCTGAGACAGGACATAAGCTGGTTTGATGACCCTAAA 2751
Qy    590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db   2752 AACACCACAGGAGCGCTGACCACCAGGCTTGCCAATGACGCTGCTCAAGTGAAAGGGGCT 2811
Qy    610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db   2812 ACAGGGTCTAGGCTTGCTGTTATTACCCAGAACATAGCAAATCTTGGGACAGGCATCATC 2871
Qy    630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db   2872 ATATCCCTGATCTACGGCTGGCAATTGACACTTTTACTCCTAGCAATTGTTCCCATCATT 2931
Qy    650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db   2932 GCTATAGCAGGAGTGGTTGAAATGAAAATGTTGTCTGGACAAGCGCTGAAAGATAAGAAG 2991
Qy    670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db   2992 GAACTAGAAGGTTCTGGGAAGATCGCTACAGAAGCAATTGAAAACTTTCGCACTGTCGTC 3051
Qy    690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db   3052 TCTTTGACTCGGGAGCAGAAGTTTGAAACTATGTATGCCCAGAGCTTGCAGATACCATAC 3111
Qy    710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db   3112 AGAAATGCTTTGAAGAAAGCGCACGTCTTTGGGATCACTTTCTCCTTCACCCAGGCCATG 3171
Qy    730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db   3172 ATGTATTTCTCCTATGCTGCTTGTTTCCGGTTTGATGCCTACTTGGTGGCACGAGAACTC 3231
Qy    750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db   3232 ATGACATTTGAAAATGTTCTGTTAGTATTCTCAGCTATTGTCTTTGGTGCCATGGCAGTG 3291
Qy    770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db   3292 GGGCAGGTCAGTTCATTCGCTCCTGACTACGCGAAAGCCAAAGTCTCGGCATCCCACATC 3351
Qy    790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db   3352 ATCAGGATCATTGAGAAAATCCCTGAGATTGACAGCTACAGCACGGAGGGCTTGAAGCCT 3411
Qy    810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db   3412 AATATGTTGGAAGGAAATGTGAAATTTAATGGAGTCATGTTCAACTATCCCACCCGACCC 3471
Qy    830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db   3472 AACATCCCAGTGCTTCAGGGGCTGAGCCTAGAGGTGAAGAAAGGGCAGACGCTGGCCCTC 3531
Qy    850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db   3532 GTGGGCAGCAGTGGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTTGAGCGCTTCTATGAC 3591
Qy    870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
Db   3592 CCCATGGCCGGAACAGTGTTTCTAGATGGCAAAGAAATAAAGCAACTCAATGTCCAGTGG 3651
Qy    890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db   3652 CTCCGCGCCCACCTGGGCATTGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATCGCC 3711
Qy    910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db   3712 GAGAACATTGCCTACGGAGACAACAGCCGTGTCGTGTCTCATAAGGAGATCGTGAAGGCA 3771
Qy    930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db   3772 GCCAAGGAGGCCAACATCCACCAGTTCATCGACTCACTGCCTGAGAAATACAACACCAGA 3831
Qy    950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db   3832 GTGGGAGACAAAGGGACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATCGCCATCGCGCGC 3891
Qy    970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db   3892 GCCCTCGTCAGACAGCCTCACATCTTACTTCTGGATGAAGCGACATCAGCTCTGGATACG 3951
Qy    990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db   3952 GAGAGTGAAAAGGTCGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACCTGCATT 4011
Qy   1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db   4012 GTGATCGCGCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATTCAGAAC 4071
Qy   1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
Db   4072 GGCCAGGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCCCAGAAAGGCATCTATTTC 4131
Qy   1050 LysLeuValAsnAlaGlnSer 1056
Db   4132 TCGATGGTCAGTGTGCAGGCT 4152
```

RESULT 9
US-09-866-866A-1
; Sequence 1, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3860
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-866A-1

```
Alignment Scores:
Pred. No.:              1.73e-280     Length:         3860
Score:                  3100.50       Matches:        595
Percent Similarity:     75.54%        Conservative:   211
Best Local Similarity:  55.76%        Mismatches:     218
Query Match:            58.17%        Indels:         43
DB:                     10            Gaps:           6
```

US-09-873-409-4 (1-1058) x US-09-866-866A-1 (1-3860)

```
Qy      4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
Db    712 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 771
Qy     24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
```

```
Db    772 TTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAAAAGAAAGAACTTGAA------ 825

Qy     44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63

Db    825 ------------------------------------------------------------ 825

Qy     64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83

Db    826 ---------------AGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 870

Qy     84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103

Db    871 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 930

Qy    104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123

Db    931 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 984

Qy    124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143

Db    985 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1044

Qy    144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163

Db   1045 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104

Qy    164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183

Db   1105 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1164

Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203

Db   1165 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1224

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223

Db   1225 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1284

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243

Db   1285 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1344

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263

Db   1345 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1404

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283

Db   1405 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1464

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303

Db   1465 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1524

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323

Db   1525 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1584

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343

Db   1585 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1644

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363

Db   1645 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1704

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383

Db   1705 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 1764

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403

Db   1765 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 1824

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423

Db   1825 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 1884

Qy    424 Gln-----------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437

Db   1885 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 1944

Qy    438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 452

Db   1945 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2004

Qy    453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470

Db   2005 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2064

Qy    471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490

Db   2065 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2124

Qy    491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510

Db   2125 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2184

Qy    511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529

Db   2185 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2244

Qy    530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549

Db   2245 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2304

Qy    550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569

Db   2305 ACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2364

Qy    570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589

Db   2365 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2424

Qy    590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609

Db   2425 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2484

Qy    610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629

Db   2485 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2544

Qy    630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649

Db   2545 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 2604

Qy    650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669

Db   2605 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 2664

Qy    670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689

Db   2665 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 2724

Qy    690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709

Db   2725 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2784

Qy    710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729

Db   2785 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 2844

Qy    730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749

Db   2845 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 2904
```

```
Qy    750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db   2905 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 2964

Qy    770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db   2965 GGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3024

Qy    790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db   3025 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3084

Qy    810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db   3085 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3144

Qy    830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db   3145 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3204

Qy    850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db   3205 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3264

Qy    870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
Db   3265 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3324

Qy    890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db   3325 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3384

Qy    910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db   3385 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3444

Qy    930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db   3445 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3504

Qy    950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db   3505 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3564

Qy    970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db   3565 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 3624

Qy    990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db   3625 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 3684

Qy   1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db   3685 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 3744

Qy   1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
Db   3745 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 3804

Qy   1050 LysLeuValAsnAlaGlnSer 1056
Db   3805 TCAATGGTCAGTGTCCAGGCT 3825
```

RESULT 10
US-09-866-866A-3
; Sequence 3, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 3860
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-866-866A-3

```
Alignment Scores:
Pred. No.:               1.73e-280    Length:         3860
Score:                   3100.50      Matches:        595
Percent Similarity:      75.54%       Conservative:   211
Best Local Similarity:   55.76%       Mismatches:     218
Query Match:             58.17%       Indels:         43
DB:                      10           Gaps:           6

US-09-873-409-4 (1-1058) x US-09-866-866A-3 (1-3860)

Qy      4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
Db    712 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 771

Qy     24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db    772 TTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAAAAGAAAGAACTTGAA------ 825

Qy     44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db    825 ------------------------------------------------------------ 825

Qy     64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db    826 --------------AGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 870

Qy     84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db    871 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 930

Qy    104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db    931 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 984

Qy    124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db    985 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1044

Qy    144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db   1045 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104

Qy    164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db   1105 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1164

Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db   1165 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1224

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db   1225 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1284

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
```

```
Db   1285 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1344

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db   1345 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1404

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db   1405 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1464

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db   1465 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1524

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db   1525 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1584

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db   1585 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1644

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db   1645 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1704

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db   1705 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 1764

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db   1765 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 1824

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   1825 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 1884

Qy    424 Gln------------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
Db   1885 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 1944

Qy    438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 452
Db   1945 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2004

Qy    453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470
Db   2005 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2064

Qy    471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
Db   2065 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2124

Qy    491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db   2125 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2184

Qy    511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
Db   2185 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2244

Qy    530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db   2245 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2304

Qy    550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db   2305 ACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2364

Qy    570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
Db   2365 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2424

Qy    590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db   2425 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2484

Qy    610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db   2485 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2544

Qy    630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db   2545 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 2604

Qy    650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db   2605 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 2664

Qy    670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db   2665 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 2724

Qy    690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db   2725 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2784

Qy    710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db   2785 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 2844

Qy    730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db   2845 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 2904

Qy    750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db   2905 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 2964

Qy    770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db   2965 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3024

Qy    790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db   3025 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3084

Qy    810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db   3085 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3144

Qy    830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db   3145 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3204

Qy    850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db   3205 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3264

Qy    870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
Db   3265 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3324

Qy    890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db   3325 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3384

Qy    910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db   3385 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3444

Qy    930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db   3445 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3504
```

```
Qy     950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
           |||||    |||||    ||||||||||||||||||||||||||:::|||||||||||
Db    3505 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3564

Qy     970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
           ||||||::::::::|||  |||||||||||||||||||||||||||||||||||
Db    3565 GCCCTTGTTAGACAGCCTCATATTTGCTTTGGATGAAGCCACGTCAGCTCTGGATACA 3624

Qy     990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
           :::||||||||||||||||||    ||||||||||||||||||    ||||||||||:::
Db    3625 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 3684

Qy    1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
           |||:::   ||||||||||    ||||||||||||||||||||||||||       |||
Db    3685 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 3744

Qy    1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
           |||::::::||||||    |||||||||||:::||||||       :::    ||||||||
Db    3745 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 3804

Qy    1050 LysLeuValAsnAlaGlnSer 1056
              :::|||:::   |||:::
Db    3805 TCAATGGTCAGTGTCCAGGCT 3825
```

RESULT 11
US-09-306-417-1
; Sequence 1, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 8630
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: proviral
;   OTHER INFORMATION: plasmid DNA
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(160)
;   OTHER INFORMATION: plasmid backbone (pUC)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (161)..(677)
;   OTHER INFORMATION: 5'-LTR
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (532)..(1219)
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: (1220)..(5062)
;   OTHER INFORMATION: m4 mdr-1 cDNA
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5215)..(5774)
;   OTHER INFORMATION: 3'-LTR
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5775)..(8630)
;   OTHER INFORMATION: plasmid backbone (pUC)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(8630)
;   OTHER INFORMATION: retroviral expression vector SFbeta71m4
US-09-306-417-1

```
Alignment Scores:
Pred. No.:               5.71e-280      Length:         8630
Score:                   3100.50        Matches:        595
Percent Similarity:      75.54%         Conservative:   211
Best Local Similarity:   55.76%         Mismatches:     218
Query Match:             58.17%         Indels:         43
DB:                      10             Gaps:           6

US-09-873-409-4 (1-1058) x US-09-306-417-1 (1-8630)

Qy       4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
           |||   |||   ||||||||   ||||||:::||||||||||||||||||||||||||||
Db    1931 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 1990

Qy      24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
           |||::::::||||||||||||||||||||         |||:::|||||||||:::
Db    1991 TTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAAAAGAAAGAACTTGAA------ 2044

Qy      44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63

Db    2044 ------------------------------------------------------------ 2044

Qy      64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
                          ||||||    :::|||||::::::||||||        ||||||||
Db    2045 ---------------AGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 2089

Qy      84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
           :::   |||   :::   :::|||:::||||||    :::       :::      :::|||
Db    2090 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 2149

Qy     104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
              ||||||||||||||||||||::::|||:::|||:::||||||       |||:::|||
Db    2150 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 2203

Qy     124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
           |||   ||||||   |||||||||||||||:::      ::::::    :::|||   |||
Db    2204 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTAGTGTTGGACAGGCA 2263

Qy     144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
              |||       |||    ||||||    |||||||||||||||:::    ||||||::::::
Db    2264 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 2323

Qy     164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
           ||||||    ||||||||||||||||||:::::::|||    :::|||:::||||||::::::|||
Db    2324 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 2383

Qy     184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
           :::|||    :::||||||:::||||||    |||:::||||||||||||        :::|||
Db    2384 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 2443

Qy     204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
           ||||||||||||||||||||||||:::::::::||||||:::||||||||||||||||||
Db    2444 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 2503

Qy     224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
           :::|||    ||||||||||||    ||||||||:::||||||||||||||||||    :::
Db    2504 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 2563

Qy     244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
           |||   :::   ||||||    ||||||       :::|||||||||      |||:::
Db    2564 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 2623

Qy     264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
              ||||||||||||||||||||||||||||||||    |||||||||:::    ||||||
Db    2624 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 2683

Qy     284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
           :::||||||||||::::::||||||    :::|||:::|||:::|||    :::|||||||||
Db    2684 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 2743

Qy     304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
           ||||||||||||||||||:::    |||:::||||||:::||||||||||||||||:::|||
```

```
Db   2744 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2803

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db   2804 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2863

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db   2864 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2923

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db   2924 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 2983

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db   2984 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 3043

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   3044 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 3103

Qy    424 Gln------------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
Db   3104 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 3163

Qy    438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 452
Db   3164 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 3223

Qy    453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470
Db   3224 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 3283

Qy    471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
Db   3284 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 3343

Qy    491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db   3344 CCTTATTTGTTGTTGGTGTATTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 3403

Qy    511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
Db   3404 GCAATAATATTTCAAAGATTATAGGGGTTTTACAAGAATTGATGATCCTGAAACAAAA 3463

Qy    530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db   3464 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 3523

Qy    550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db   3524 ACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 3583

Qy    570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
Db   3584 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 3643

Qy    590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db   3644 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 3703

Qy    610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db   3704 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 3763

Qy    630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db   3764 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 3823

Qy    650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db   3824 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 3883

Qy    670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db   3884 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 3943

Qy    690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db   3944 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 4003

Qy    710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db   4004 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 4063

Qy    730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db   4064 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 4123

Qy    750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db   4124 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 4183

Qy    770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db   4184 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 4243

Qy    790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db   4244 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 4303

Qy    810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db   4304 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 4363

Qy    830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db   4364 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 4423

Qy    850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db   4424 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 4483

Qy    870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
Db   4484 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATTAAGCGACTGAATGTTCAGTGG 4543

Qy    890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db   4544 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 4603

Qy    910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db   4604 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 4663

Qy    930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db   4664 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 4723

Qy    950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db   4724 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 4783

Qy    970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db   4784 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 4843

Qy    990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db   4844 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 4903

Qy   1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db   4904 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 4963
```

```
Qy    1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
Db    4964 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 5023

Qy    1050 LysLeuValAsnAlaGlnSer 1056
Db    5024 TCAATGGTCAGTGTCCAGGCT 5044

RESULT 12
US-09-306-417-2
; Sequence 2, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
;  APPLICANT: Heinrich-Pette-Institut
;  TITLE OF INVENTION: Retroviral Gene Transfer Vectors
;  FILE REFERENCE: P50491
;  CURRENT APPLICATION NUMBER: US/09/306,417
;  CURRENT FILING DATE: 1999-05-06
;  EARLIER APPLICATION NUMBER: DE 198 22 115
;  EARLIER FILING DATE: 1998-05-08
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 8630
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: proviral
;   OTHER INFORMATION: plasmid DNA
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(8630)
;   OTHER INFORMATION: retroviral expression vector SFbeta91mSA1
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(160)
;   OTHER INFORMATION: plasmid backbone (pUC)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (161)..(677)
;   OTHER INFORMATION: 5'-LTR
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (532)..(1219)
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: (1220)..(5062)
;   OTHER INFORMATION: mSA1 mdr1 cDNA
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5215)..(5774)
;   OTHER INFORMATION: 3'-LTR
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5775)..(8630)
;   OTHER INFORMATION: plasmid backbone (pUC)
US-09-306-417-2

Alignment Scores:
Pred. No.:             5.71e-280      Length:        8630
Score:                 3100.50        Matches:       595
Percent Similarity:    75.54%         Conservative:  211
Best Local Similarity: 55.76%         Mismatches:    218
Query Match:           58.17%         Indels:        43
DB:                    10             Gaps:          6

US-09-873-409-4 (1-1058) x US-09-306-417-2 (1-8630)

Qy       4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
Db    1931 TCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTC 1990

Qy      24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db    1991 TTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAAAAGAAAGAACTTGAA------ 2044

Qy      44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db    2044 ------------------------------------------------------------ 2044

Qy      64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db    2045 ---------------AGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 2089

Qy      84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db    2090 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 2149

Qy     104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db    2150 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 2203

Qy     124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db    2204 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 2263

Qy     144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db    2264 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 2323

Qy     164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db    2324 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 2383

Qy     184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db    2384 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 2443

Qy     204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db    2444 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 2503

Qy     224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db    2504 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 2563

Qy     244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db    2564 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 2623

Qy     264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db    2624 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 2683

Qy     284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db    2684 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 2743

Qy     304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db    2744 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2803

Qy     324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db    2804 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2863

Qy     344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db    2864 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2923

Qy     364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db    2924 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 2983

Qy     384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
```

```
Db   2984 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 3043
Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   3044 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 3103
Qy    424 Gln-----------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 437
Db   3104 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 3163
Qy    438 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 452
Db   3164 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 3223
Qy    453 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 470
Db   3224 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 3283
Qy    471 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 490
Db   3284 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 3343
Qy    491 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 510
Db   3344 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 3403
Qy    511 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 529
Db   3404 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 3463
Qy    530 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 549
Db   3464 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 3523
Qy    550 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 569
Db   3524 ACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 3583
Qy    570 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 589
Db   3584 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 3643
Qy    590 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 609
Db   3644 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 3703
Qy    610 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 629
Db   3704 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 3763
Qy    630 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 649
Db   3764 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 3823
Qy    650 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 669
Db   3824 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 3883
Qy    670 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 689
Db   3884 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 3943
Qy    690 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 709
Db   3944 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 4003
Qy    710 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 729
Db   4004 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 4063
Qy    730 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 749
Db   4064 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 4123
Qy    750 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 769
Db   4124 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 4183
Qy    770 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 789
Db   4184 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 4243
Qy    790 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 809
Db   4244 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 4303
Qy    810 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 829
Db   4304 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 4363
Qy    830 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 849
Db   4364 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 4423
Qy    850 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 869
Db   4424 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 4483
Qy    870 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 889
Db   4484 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 4543
Qy    890 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 909
Db   4544 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 4603
Qy    910 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 929
Db   4604 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 4663
Qy    930 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 949
Db   4664 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 4723
Qy    950 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 969
Db   4724 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 4783
Qy    970 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 989
Db   4784 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 4843
Qy    990 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1009
Db   4844 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 4903
Qy   1010 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1029
Db   4904 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 4963
Qy   1030 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1049
Db   4964 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 5023
Qy   1050 LysLeuValAsnAlaGlnSer 1056
Db   5024 TCAATGGTCAGTGTCCAGGCT 5044
```

RESULT 13
US-09-866-866A-7
; Sequence 7, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells

```
;  FILE REFERENCE: 1340-1-021CIP2
;  CURRENT APPLICATION NUMBER: US/09/866,866A
;  CURRENT FILING DATE:  2001-08-30
;  PRIOR APPLICATION NUMBER: 09/584,586
;  PRIOR FILING DATE: 2000-05-31
;  PRIOR APPLICATION NUMBER: PCT/US99/11825
;  PRIOR FILING DATE: 1999-05-27
;  PRIOR APPLICATION NUMBER: 60/086,988
;  PRIOR FILING DATE: 1998-05-28
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
;   LENGTH: 4788
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-866-866A-7

Alignment Scores:
Pred. No.:              2.06e-279     Length:         4788
Score:                  3090.50       Matches:        588
Percent Similarity:     75.12%        Conservative:   215
Best Local Similarity:  55.00%        Mismatches:     219
Query Match:            57.98%        Indels:         47
DB:                     10            Gaps:           6

US-09-873-409-4 (1-1058) x US-09-866-866A-7 (1-4788)

Qy      4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
Db    700 TCATTTACTGATAAGGAACTCCATGCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTC 759

Qy     24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db    760 TTAGCAGCCATCAGAACTGTGATTGCGTTTGGAGGACAAAAGAAGGAACTTGAA------ 813

Qy     44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db    813 ------------------------------------------------------------ 813

Qy     64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db    814 ---------------AGGTACAATAACAACTTGGAAGAAGCTAAAAGGCTGGGGATAAAG 858

Qy     84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db    859 AAAGCTATCACGGCCAACATCTCCATGGGTGCAGCTTTTCTCCTTATCTATGCATCATAT 918

Qy    104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db    919 GCTCTGGCATTCTGGTATGGACTTCCTTGGTCATCTCCAAAGAA------TACTCTATT 972

Qy    124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db    973 GGACAAGTGCTCACTGTCTTCTTTTCCGTGTTAATTGGAGCATTCAGTGTTGGACAGGCA 1032

Qy    144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db   1033 TCTCCAAATATTGAAGCCTTCGCCAATGCACGAGGAGCAGCTTATGAAGTCTTCAAAATA 1092

Qy    164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db   1093 ATTGATAATAAGCCCAGTATAGACAGCTTCTCAAAGAGTGGGCACAAACCAGACAACATA 1152

Qy    184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db   1153 CAAGGAAATCTGGAATTTAAGAATATTCACTTCAGTTACCCATCTCGAAAAGAAGTTCAG 1212

Qy    204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db   1213 ATCTTGAAGGGCCTCAATCTGAAGGTGAAGAGCGGACAGACGGTGGCCCTGGTTGGCAAC 1272

Qy    224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db   1273 AGTGGCTGTGGAAAAAGCACAACTGTCCAGCTGATGCAAAGGCTCTACGACCCCCTAGAT 1332

Qy    244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db   1333 GGCATGGTCAGTATCGACGGACAGGACATCAGAACCATCAATGTGAGGTATCTGAGGGAG 1392

Qy    264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db   1393 ATCATTGGTGTGGTGAGTCAGGAACCTGTGCTGTTTGCCACCACGATCGCCGAGAACATT 1452

Qy    284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db   1453 CGCTATGGCCGAGAAGATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1512

Qy    304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db   1513 GCCTATGACTTCATCATGAAACTGCCCCACCAATTTGACACCCTGGTTGGTGAGAGAGGG 1572

Qy    324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db   1573 GCGCACGTGAGTGGGGGACAGAAACAGAGAATCGCCATTGCCCGGGCCCTGGTCCGCAAT 1632

Qy    344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db   1633 CCCAAGATCCTTTTGTTGGACGAGGCCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTG 1692

Qy    364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db   1693 GTTCAGGCCGCACTGGATAAGGCTAGAGAAGGCCGGACCACCATTGTGATAGCTCATCGC 1752

Qy    384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db   1753 TTGTCTACCGTTCGTAATGCTGACGTCATTGCTGGTTTTGATGGTGGTGTCATTGTGGAG 1812

Qy    404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 423
Db   1813 CAAGGAAATCATGATGAGCTCATGAGAGAAAAGGGCATTTACTTCAAACTTGTCATGACA 1872

Qy    424 Gln------------------------------AspIleLysLysAlaAspGluGlnMet 433
Db   1873 CAGACAGCAGGAAATGAAATTGAATTAGGAAATGAAGCTTGTAAATCTAAGGATGAAATT 1932

Qy    434 GluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLys 453
Db   1933 GATAATTTAGACATGTCTTCAAAAGATTCAGGATCCAGTCTAATAAGAAGA------AGA 1986

Qy    454 SerIleLysSerAspPheIle---------------AspLysAlaGluGluSerThrGln 468
Db   1987 TCAACTCGCAAAAGCATCTGTGGACCACATGACCAAGACAGGAAGCTTAGTACCAAAGAG 2046

Qy    469 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 488
Db   2047 GCCCTGGATGAAGATGTACCTCCAGCTTCCTTTTGGCGGATCCTGAAGTTGAATTCAACT 2106

Qy    489 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 508
Db   2107 GAATGGCCTTATTTTGTGGTTGGTATATTCTGTGCCATAATAAATGGAGGCTTACAGCCA 2166

Qy    509 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThr 527
Db   2167 GCATTCTCCGTAATATTTTCAAAAGTTGTAGGGGTTTTTACAAATGGTGGCCCCCCTGAA 2226

Qy    528 ThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 547
Db   2227 ACCCAGCGGCAGAACAGCAACTTGTTTTCCTTGTTGTTTCTGATCCTTGGGATCATTTCT 2286

Qy    548 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 567
Db   2287 TTCATTACATTTTTTCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAG 2346

Qy    568 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 587
Db   2347 CGACTCCGATACATGGTTTTCAAATCCATGCTGAGACAGGATGTGAGCTGGTTTGATGAC 2406
```

```
Qy     588 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 607
Db    2407 CCTAAAAACACCACCGGAGCACTGACCACCAGGCTCGCCAACGATGCTGCTCAAGTGAAA 2466

Qy     608 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 627
Db    2467 GGGGCTACAGGGTCTAGGCTTGCTGTGATTTTCCAGAACATAGCAAATCTTGGGACAGGA 2526

Qy     628 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 647
Db    2527 ATCATCATATCCCTAATCTATGGCTGGCAACTAACACTTTTACTCTTAGCAATTGTACCC 2586

Qy     648 ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp 667
Db    2587 ATCATTGCGATAGCTGGAGTGGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGAT 2646

Qy     668 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 687
Db    2647 AAGAAGGAACTAGAAGGTTCTGGAAAGATTGCTACGGAAGCAATTGAAAACTTCCGCACT 2706

Qy     688 IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 707
Db    2707 GTTGTCTCTTTGACTCGGGAGCAGAAGTTTGAAACCATGTATGCCCAGAGCTTGCAGATA 2766

Qy     708 GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis 727
Db    2767 CCATACAGAAATGCGATGAAGAAAGCACACGTGTTTGGGATCACGTTCTCCTTCACCCAG 2826

Qy     728 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla 747
Db    2827 GCCATGATGTATTTTTCTTATGCTGCTTGTTTCCGGTTCGGTGCCTACTTGGTGACACAA 2886

Qy     748 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 767
Db    2887 CAACTCATGACTTTTGAAAATGTTCTGTTAGTATTCTCAGCTATTGTCTTTGGTGCCATG 2946

Qy     768 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 787
Db    2947 GCAGTGGGGCAGGTCAGTTCATTCGCTCCTGACTATGCGAAAGCAACAGTGTCAGCATCC 3006

Qy     788 HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys 807
Db    3007 CACATCATCAGGATCATTGAGAAAACCCCCGAGATTGACAGCTACAGCACGCAAGGCCTA 3066

Qy     808 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys 827
Db    3067 AAGCCGAATATGTTGGAAGGAAATGTGCAATTTAGTGGAGTCGTGTTCAACTATCCCACC 3126

Qy     828 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 847
Db    3127 CGACCCAGCATCCCAGTGCTTCAGGGGCTGAGCCTTGAGGTGAAGAAGGGCCAGACGCTG 3186

Qy     848 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 867
Db    3187 GCCCTGGTGGGCAGCAGTGGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAGCGCTTC 3246

Qy     868 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 887
Db    3247 TACGACCCCATGGCTGGATCAGTGTTTCTAGATGGCAAAGAAATAAAGCAACTGAATGTC 3306

Qy     888 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 907
Db    3307 CAGTGGCTCCGAGCACAGCTGGGCATTGTGTCCCAAGAGCCCATTCTCTTTGACTGCAGC 3366

Qy     908 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 927
Db    3367 ATCGCAGAGAACATTGCCTACGGAGACAACAGCCGGGTCGTGTCTTATGAGGAGATTGTG 3426

Qy     928 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn 947
Db    3427 AGGGCAGCCAAGGAGGCCAACATCCACCAGTTCATCGACTCGCTACCTGATAAATACAAC 3486

Qy     948 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 967
Db    3487 ACCAGAGTAGGAGACAAAGGCACTCAGCTGTCGGGTGGGCAGAAGCAGCGCATCGCCATC 3546

Qy     968 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 987
Db    3547 GCACGCGCCCTCGTCAGACAGCCTCACATTTTACTTCTGGACGAAGCAACATCAGCTCTG 3606

Qy     988 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1007
Db    3607 GATACAGAAAGTGAAAAGGTTGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACC 3666

Qy    1008 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1027
Db    3667 TGCATTGTGATCGCTCACCGCCTGTCCACCATCCAGAACGCGGACTTGATCGTGGTGATT 3726

Qy    1028 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1047
Db    3727 CAGAACGGCAAGGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCGCAGAAGGGCATC 3786

Qy    1048 TyrPheLysLeuValAsnAlaGlnSer 1056
Db    3787 TACTTCTCAATGGTCAGTGTGCAGGCT 3813
```

RESULT 14
US-09-880-107-2299
; Sequence 2299, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2299
;   LENGTH: 3924
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 M23234
US-09-880-107-2299

Alignment Scores:

Pred. No.:	5.02e-279	Length:	3924
Score:	3085.00	Matches:	597
Percent Similarity:	74.72%	Conservative:	198
Best Local Similarity:	56.11%	Mismatches:	229
Query Match:	57.88%	Indels:	40
DB:	10	Gaps:	5

US-09-873-409-4 (1-1058) x US-09-880-107-2299 (1-3924)

```
Qy       4 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 23
Db     750 GCATTTAGTGACAAAGAACTAGCTGCTTATGCAAAAGCAGGCGCCGTGGCAGAAGAGGCT 809

Qy      24 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 43
Db     810 CTGGGGGCCATCAGGACTGTGATAGCTTTCGGGGGCCAGAACAAAGAGCTGGAA------ 863

Qy      44 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 63
Db     863 ------------------------------------------------------------ 863
```

```
Qy    64 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 83
Db   864 ---------------AGGTATCAGAAACATTTAGAAAATGCCAAAGAGATTGGAATTAAA 908

Qy    84 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 103
Db   909 AAAGCTATTTCAGCAAACATTTCCATGGGTATTGCCTTCCTGTTAATATATGCATCATAT 968

Qy   104 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 123
Db   969 GCACTGGCCTTCTGGTATGGATCCACTCTAGTCATATCAAAAGAA------TATACTATT 1022

Qy   124 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 143
Db  1023 GGAAATGCAATGACAGTTTTTTTTCAATCCTAATTGGAGCTTTCAGTGTTGGCCAGGCT 1082

Qy   144 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 163
Db  1083 GCCCCATGTATTGATGCTTTTGCCAATGCAAGAGGAGCAGCATATGTGATCTTTGATATT 1142

Qy   164 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 183
Db  1143 ATTGATAATAATCCTAAAATTGACAGTTTTTCAGAGAGAGGACACAAACCAGACAGCATC 1202

Qy   184 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 203
Db  1203 AAAGGGAATTTGGAGTTCAATGATGTTCACTTTTCTTACCCTTCTCGAGCTAACGTCAAG 1262

Qy   204 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 223
Db  1263 ATCTTGAAGGGCCTCAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAGT 1322

Qy   224 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 243
Db  1323 AGTGGCTGTGGGAAGAGCACAACGGTCCAGCTGATACAGAGGCTCTATGACCCTGATGAG 1382

Qy   244 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 263
Db  1383 GGCACAATTAACATTGATGGGCAGGATATTAGGAACTTTAATGTAAACTATCTGAGGGAA 1442

Qy   264 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 283
Db  1443 ATCATTGGTGTGGTGAGTCAGGAGCCGGTGCTGTTTTCCACCACAATTGCTGAAAATATT 1502

Qy   284 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 303
Db  1503 TGTTATGGCCGTGGAAATGTAACCATGGATGAGATAAAGAAAGCTGTCAAAGAGGCCAAC 1562

Qy   304 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 323
Db  1563 GCCTATGAGTTTATCATGAAATTACCACAGAAATTTGACACCCTGGTTGGAGAGAGAGGG 1622

Qy   324 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 343
Db  1623 GCCCAGCTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1682

Qy   344 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 363
Db  1683 CCCAAGATCCTTCTGCTGGATGAGGCCACGTCAGCATTGGACACAGAAAGTGAAGCTGAG 1742

Qy   364 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 383
Db  1743 GTACAGGCAGCTCTGGATAAGGCCAGAGAAGGCCGGACCACCATTGTGATAGCACACCGA 1802

Qy   384 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 403
Db  1803 CTGTCTACGGTCCGAAATGCAGATGTCATCGCTGGGTTTGAGGATGGAGTAATTGTGGAG 1862

Qy   404 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal------ 421
Db  1863 CAAGGAAGCCACAGCGAACTGATGAAGAAGGAAGGGGTGTACTTCAAACTTGTCAACATG 1922

Qy   422 --------------------MetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 434
Db  1923 CAGACATCAGGAAGCCAGATCCAGTCAGAAGAATTTGAACTAAATGATGAAAAGGCTGCC 1982

Qy   435 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 454
Db  1983 ACTAGAATGGCCCCAAATGGCTGGAAATCTCGCCTATTTAGGCATTCTACTCAGAAAAAC 2042

Qy   455 IleLys---SerAspPheIleAspLysAla---GluGluSerThrGlnSerLysGluIle 472
Db  2043 CTTAAAAATTCACAAATGTGTCAGAAGAGCCTTGATGTGGAAACCGATGGACTTGAAGCA 2102

Qy   473 SerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPhe 492
Db  2103 AATGTGCCACCAGTGTCCTTTCTGAAGGTCCTGAAACTGAATAAAACAGAATGGCCCTAC 2162

Qy   493 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 512
Db  2163 TTTGTCGTGGGAACAGTATGTGCCATTGCCAATGGGGGGCTTCAGCCGGCATTTTCAGTC 2222

Qy   513 IlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAsp 532
Db  2223 ATATTCTCAGAGATCATAGCGATTTTTGGACCAGGCGATGATGCAGTGAAGCAGCAGAAG 2282

Qy   533 AlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPhe 552
Db  2283 TGCAACATATTCTCTTTGATTTTCTTATTTCTGGGAATTATTTCTTTTTTTACTTTCTTC 2342

Qy   553 MetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeu 572
Db  2343 CTTCAGGGTTTCACGTTTGGGAAAGCTGGCGAGATCCTCACCAGAAGACTGCGGTCAATG 2402

Qy   573 AlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThr 592
Db  2403 GCTTTTAAAGCAATGCTAAGACAGGACATGAGCTGGTTTGATGACCATAAAAACAGTACT 2462

Qy   593 GlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySer 612
Db  2463 GGTGCACTTTCTACAAGACTTGCCACAGATGCTGCCCAAGTCCAAGGAGCCACAGGAACC 2522

Qy   613 ArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPhe 632
Db  2523 AGGTTGGCTTTAATTGCACAGAATATAGCTAACCTTGGAACTGGTATTATCATATCATTT 2582

Qy   633 IleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThr 652
Db  2583 ATCTACGGTTGGCAGTTAACCCTATTGCTATTAGCAGTTGTTCCAATTATTGCTGTGTCA 2642

Qy   653 GlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLys 672
Db  2643 GGAATTGTTGAAATGAAATTGTTGGCTGGAAATGCCAAAAGAGATAAAAAAGAACTGGAA 2702

Qy   673 HisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThr 692
Db  2703 GCTGCTGGAAAGATTGCAACAGAGGCAATAGAAAATATTAGGACAGTTGTGTCTTTGACC 2762

Qy   693 ArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThr 712
Db  2763 CAGGAAAGAAAATTTGAATCAATGTATGTTGAAAAATTGTATGGACCTTACAGGAATTCT 2822

Qy   713 SerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPhe 732
Db  2823 GTGCAGAAGGCACACATCTATGGAATTACTTTTAGTATCTCACAAGCATTTATGTATTTT 2882

Qy   733 AlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrPro 752
Db  2883 TCCTATGCCGGTTGTTTTCGATTTGGTGCATATCTCATTGTGAATGGACATATGCGCTTC 2942

Qy   753 GluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThr 772
Db  2943 AGAGATGTTATTCTGGTGTTTTCTGCAATTGTATTTGGTGCAGTGGCTCTAGGACATGCC 3002

Qy   773 LeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeu 792
```

```
Db    3003 AGTTCATTTGCTCCAGACTATGCTAAAGCTAAGCTGTCTGCAGCCCACTTATTCATGCTG 3062

Qy     793 LeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCys 812
                 |||::::::|||   ||||||||   |||:::|||||   ||||||||
Db    3063 TTTGAAAGACAACCTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAGCCTGATAAATTT 3122

Qy     813 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 832
           |||||||||:::   |||   ||||||   |||   ||||||   |||   :::|||
Db    3123 GAAGGAAATATAACATTTAATGAAGTCGTGTTCAACTATCCCACCCGAGCAAACGTGCCA 3182

Qy     833 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 852
           :::|||:::||||||||||||   :::::::::|||:::|||:::|||   ||||||||
Db    3183 GTGCTTCAGGGGCTGAGCCTGGAGGTGAAGAAAGGCCAGACACTAGCCCTGGTGGGCAGC 3242

Qy     853 SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGln 872
           ||||||||||||||||||   ||||||||||||:::|||   |||||||||:::
Db    3243 AGTGGCTGTGGGAAGAGCACGGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCG 3302

Qy     873 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 892
           |||   ||||||   ||||||   :::||||||:::|||||||||||||||||||||:::
Db    3303 GGGACAGTGCTTCTCGATGGTCAAGAAGCAAAGAAACTCAATGTCCAGTGGCTCAGAGCT 3362

Qy     893 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 912
           |||:::   ||||||   |||||||||:::||||||:::||||||||||||||||||||
Db    3363 CAACTCGGAATCGTGTCTCAGGAGCCTATCCTATTTGACTGCAGCATTGCCGAGAATATT 3422

Qy     913 AlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla 932
           ||||||||||||||||||||||||||||||   |||||||||   ||||||   |||
Db    3423 GCCTATGGAGACAACAGCCGGGTTGTATCACAGGATGAAATTGTGAGTGCAGCCAAAGCT 3482

Qy     933 AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeu 952
           ||||||||||||   |||||||||   ||||||   ||||||   |||:::||||||
Db    3483 GCCAACATACATCCTTTCATCGAGACGTTACCCCACAAATATGAAACAAGAGTGGGAGAT 3542

Qy     953 LysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeu 972
           ||||||   ||||||||||||||||||||||||||||:::||||||||||||||||||:::
Db    3543 AAGGGGACTCAGCTCTCAGGAGGTCAAAAACAGAGGATTGCTATTGCCCGAGCCCTCATC 3602

Qy     973 GlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGlu 992
           ::::::|||:::|||||||||||||||||||||||||||||||||   :::||||||
Db    3603 AGACAACCTCAAATCCTCCTGTTGGATGAAGCTACATCAGCTCTGGATACTGAAAGTGAA 3662

Qy     993 LysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThr 1012
           ||||||||||||   ||||||||||||||||||   ||||||||||||:::|||:::
Db    3663 AAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCT 3722

Qy    1013 HisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIle 1032
           ||||||||||||   |||||||||||||||||||||||||||   ||||||::::::
Db    3723 CACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGGAGAGTC 3782

Qy    1033 LysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuVal 1052
           ||||||   ||||||||||||:::||||||   :::   ||||||||   :::|||
Db    3783 AAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTC 3842

Qy    1053 AsnAlaGlnSer 1056
           :::   |||:::
Db    3843 AGTGTCCAGGCT 3854

RESULT 15
US-09-866-866A-5
; Sequence 5, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
;  APPLICANT: Sorrentino, Brian
;  APPLICANT:  Schuetz, John
;  TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
;  FILE REFERENCE: 1340-1-021CIP2
;  CURRENT APPLICATION NUMBER: US/09/866,866A
;  CURRENT FILING DATE:  2001-08-30
;  PRIOR APPLICATION NUMBER: 09/584,586
;  PRIOR FILING DATE: 2000-05-31
;  PRIOR APPLICATION NUMBER: PCT/US99/11825
;  PRIOR FILING DATE: 1999-05-27
;  PRIOR APPLICATION NUMBER: 60/086,988
;  PRIOR FILING DATE: 1998-05-28
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 4189
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-866-866A-5

Alignment Scores:
Pred. No.:                 5.53e-279       Length:         4189
Score:                     3085.00         Matches:        595
Percent Similarity:        74.93%          Conservative:   206
Best Local Similarity:     55.66%          Mismatches:     222
Query Match:               57.88%          Indels:         46
DB:                        10              Gaps:           7

US-09-873-409-4 (1-1058) x US-09-866-866A-5 (1-4189)

Qy       1 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 20
           ::::::   |||   |||:::|||||||||   ||||||:::||||||||||||||||||
Db     700 GTATTGACTTCATTTACTAATAAGGAACTCCAGGCTTATGCAAAAGCTGGAGCAGTTGCT 759

Qy      21 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 40
           ||||||||||||::::::||||||||||||||||||   |||:::||||||||
Db     760 GAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCCTTTGGAGGACAACAGAAGGAACTT 819

Qy      41 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 60
           :::
Db     820 GAA--------------------------------------------------------- 822

Qy      61 LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 80
                                    ||||||   :::||||||::::::||||||:::
Db     823 -----------------------AGGTACAATAAAAATTTAGAAGAAGCTAAAAATGTT 858

Qy      81 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 100
           |||||||||:::   |||   :::   :::|||:::|||   |||      :::
Db     859 GGCATAAAGAAAGCTATCACAGCCAGCATTTCGATAGGCATTGCCTACCTGTTGGTCTAT 918

Qy     101 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 120
              :::|||   |||||||||||||||||||||||||:::|||:::   |||
Db     919 GCATCATATGCACTGGCATTCTGGTATGGGACATCCTTGGTCCTCTCAAATGAA------ 972

Qy     121 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 140
           |||:::||||||   ||||||   ||||||||||||::::::      ::::::   |||
Db     973 TATTCTATTGGAGAAGTGCTTACTGTCTTCTTCTCTATTTTGTTGGGGACTTTTAGTATT 1032

Qy     141 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 160
           |||      |||:::   |||   ||||||   ||||||||||||||||||   |||
Db    1033 GGACACTTGGCCCCAAACATAGAAGCCTTTGCAAACGCACGAGGGGCAGCCTTTGAAATC 1092

Qy     161 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 180
           |||::::::||||||   :::||||||||||||:::|||   ||||||||||||
Db    1093 TTCAAGATAATTGATAACGAGCCAAGCATTGACAGCTTCTCAACAAAGGGCTACAAACCA 1152

Qy     181 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 200
           :::||||||   |||   :::|||||||||||||||   ||||||||||||||||||
Db    1153 GACAGTATAATGGGAAACTTAGAGTTTAAAAATGTTCACTTCAACTACCCATCGAGAAGC 1212

Qy     201 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 220
              ::::::|||||||||||||||||||||::::::|||||||||:::||||||||||||
Db    1213 GAAGTTCAGATCTTGAAGGGCCTCAATCTGAAGGTGAAGAGCGGACAGACGGTGGCCTTG 1272

Qy     221 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 240
           ||||||   :::|||   ||||||||||||   |||||||||:::||||||||||||||||
Db    1273 GTTGGCAACAGTGGCTGTGGAAAAAGCACAACTGTCCAGCTGATGCAGAGGCTCTACGAC 1332

Qy     241 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 260
           |||   :::|||   :::   :::|||   |||||||||   :::|||||||||:::
Db    1333 CCCCTGGAGGGCGTGGTCAGTATCGACGGACAAGACATCAGAACCATCAATGTGAGGTAT 1392

Qy     261 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 280
```

```
Db   1393 CTGAGGGAGATCATTGGTGTGGTGAGTCAGGAACCTGTGCTGTTTGCCACCACGATCGCC 1452

Qy    281 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 300
Db   1453 GAGAACATTCGCTATGGCCGAGAAGATGTCACCATGGATGAGATTGAGAAAGCTGTCAAG 1512

Qy    301 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 320
Db   1513 GAAGCCAATGCCTATGACTTCATCATGAAACTGCCCCACCAATTTGACACCCTGGTTGGT 1572

Qy    321 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 340
Db   1573 GAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGAATCGCCATTGCCCGGGCCCTG 1632

Qy    341 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 360
Db   1633 GTCCGCAATCCCAAGATCCTTTTGTTGGACGAGGCCACCTCAGCCCTGGATACAGAAAGT 1692

Qy    361 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 380
Db   1693 GAAGCTGTGGTGCAGGCCGCACTGGATAAGGCTAGAGAAGGCCGGACCACCATTGTGATA 1752

Qy    381 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 400
Db   1753 GCTCATCGCTTGTCTACAGTTCGTAATGCTGACGTCATTGCTGGTTTTGATGGTGGTGTC 1812

Qy    401 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 420
Db   1813 ATTGTGGAGCAAGGAAATCATGATGAGCTCATGAGAGAAAAGGGCATTTACTTCAAACTT 1872

Qy    421 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 440
Db   1873 GTCATGACACAG---ACTAGAGGAAATGAAATTGAACCAGGAAATAATGCTTATGGATCC 1929

Qy    441 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 460
Db   1930 CAGAGTGACACTGATGCTTCTGAACTGACTTCAGAAGAATCC---AAATCACCTTTAATA 1986

Qy    461 AspLysAla------------------GluGluSerThrGlnSerLysGluIleSerLeu 474
Db   1987 AGGAGATCAATTTACAGAAGTGTCCACAGAAAGCAAGACCAAGAGAGAAGACTTAGTATG 2046

Qy    475 ------------------------ProGluValSerLeuLeuLysIleLeuLysLeuAsn 486
Db   2047 AAAGAGGCTGTGGATGAAGATGTGCCTCTGGTTTCCTTTTGGCGGATCCTAAATCTAAAT 2106

Qy    487 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 506
Db   2107 CTAAGTGAATGGCCTTATTTACTTGTTGGCGTACTTTGCGCTGTTATAAATGGGTGCATA 2166

Qy    507 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGly---AsnAsnAsp 525
Db   2167 CAACCAGTGTTTGCCATAGTATTTTCAAGGATTGTAGGGGTTTTTTCAAGAGATGATGAC 2226

Qy    526 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 545
Db   2227 CATGAAACTAAACGACAGAATTGTAATTTGTTTTCCCTGTTCTTTCTGGTTATGGGGCTG 2286

Qy    546 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 565
Db   2287 ATTTCTTTTGTTACATATTTCTTTCAGGGCTTCACATTTGGCAAAGCCGGAGAGATCCTC 2346

Qy    566 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 585
Db   2347 ACCAAGCGAGTCCGATACATGGTTTTCAAATCCATGCTGAGACAGGATATAAGCTGGTTC 2406

Qy    586 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 605
Db   2407 GATGACCATAAGAACAGCACTGGCTCACTGACCACCAGGCTCGCCAGTGATGCTTCTAGT 2466

Qy    606 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 625
Db   2467 GTTAAAGGGGCGATGGGCGCCAGGCTTGCTGTAGTTACCCAGAATGTAGCAAACCTCGGG 2526

Qy    626 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 645
Db   2527 ACAGGAGTCATCCTCTCCTTAGTCTATGGCTGGCAGCTGACACTTCTACTTGTAGTAATT 2586

Qy    646 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 665
Db   2587 ATACCGCTCATTGTATTGGGCGGAATTATTGAAATGAAGCTGTTGTCTGGCCAAGCCTTG 2646

Qy    666 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 685
Db   2647 AAGGACAAGAAACAGCTTGAGATCTCTGGGAAGATTGCTACAGAAGCAATTGAAAACTTC 2706

Qy    686 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 705
Db   2707 CGCACTATTGTCTCTTTGACTCGGGAGCAGAAGTTTGAAACCATGTATGCCCAGAGCTTG 2766

Qy    706 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 725
Db   2767 CAGGTACCATACAGAAATGCGATGAAGAAAGCACACGTGTTTGGGATCACGTTCTCCTTC 2826

Qy    726 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 745
Db   2827 ACCCAGGCCATGATGTATTTTTCTTATGCTGCTTGTTTCCGGTTCGGTGCCTACTTGGTG 2886

Qy    746 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 765
Db   2887 GCACAACAACTCATGACTTTTGAAAATGTTATGTTGGTATTTTCTGCTGTTGTCTTTGGT 2946

Qy    766 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 785
Db   2947 GCCATGGCAGCTGGGAATACTAGTTCATTTGCTCCTGACTATGCGAAAGCCAAAGTATCA 3006

Qy    786 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 805
Db   3007 GCATCTCATATCATCAGGATCATTGAGAAAACCCCTGAGATTGACAGCTACAGCACAGAG 3066

Qy    806 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 825
Db   3067 GGCTTGAAGCCTACTCTGTTAGAAGGAAATGTAAAATTTAATGGAGTCCAGTTTAACTAT 3126

Qy    826 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 845
Db   3127 CCCACCCGACCCAACATCCCAGTGCTTCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAG 3186

Qy    846 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 865
Db   3187 ACGTTGGCCCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTGCTCGAG 3246

Qy    866 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 885
Db   3247 CGCTTCTACGACCCCATGGCTGGATCAGTGTTTCTAGATGGCAAAGAAATAAAGCAACTG 3306

Qy    886 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 905
Db   3307 AATGTCCAGTGGCTCCGAGCTCACCTTGGCATTGTGTCCCAGGAGCCCATTCTCTTTGAC 3366

Qy    906 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 925
Db   3367 TGCAGCATTGCAGAGAACATCGCCTATGGAGACAACAGCCGGGCCGTGTCTCATGAGGAG 3426

Qy    926 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 945
Db   3427 ATTGTGAGGGCAGCCAAGGAGGCCAACATCCACCAGTTCATCGACTCACTGCCTGATAAA 3486

Qy    946 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 965
Db   3487 TACAACACCAGAGTAGGAGACAAAGGCACTCAGCTGTCGGGTGGGCAGAAGCAGCGCATC 3546

Qy    966 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 985
Db   3547 GCCATCGCACGTGCCCTCGTCAGACAGCCTCACATTTTACTTCTGGACGAAGCAACATCA 3606
```

```
Qy      986 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1005
            |||||||||   :::||||||||||||||||||   ||||||||||||||||||   |||
Db     3607 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGC 3666

Qy     1006 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1025
            |||||||||:::|||:::   ||||||||||||   ||||||||||||||||||||||||
Db     3667 CGCACCTGCATTGTGATCGCTCACCGCCTGTCCACCATCCAGAACGCGGACTTGATCGTG 3726

Qy     1026 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1045
            |||:::   |||||||||:::||||||   ||||||||||||:::||||||      :::
Db     3727 GTGATTGAGAACGGCAAAGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCGCAGAAG 3786

Qy     1046 AspIleTyrPheLysLeuValAsnAla 1054
               |||||||||   :::|||   |||
Db     3787 GGCATCTACTTCTCAATGGTCCAGGCT 3813
```

Search completed: March 31, 2003, 16:04:42
Job time : 638.104 secs

THIS PAGE BLANK (USPTO)

```
                 GenCore version 5.1.4_p5_4578
          Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:        March 31, 2003, 13:52:37 ; Search time 444.798 Seconds
                                          (without alignments)
                                          2337.888 Million cell updates/sec

Title:          US-09-873-409-5
Perfect score:  6159
Sequence:       1 MILGILASLVNGACLPLMPL..........QELLRNRDIYFKLVNAQSVQ 1222

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters:      1148742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115424_19290/app_query.fasta_1.7544
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09873409_@CGN_1_1_960_@runat_27032003_115424_19290
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Published_Applications_NA:*
                1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
                2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
                3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
                4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
                5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
                6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
                7:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
                8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
                9:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
                10:  /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
                11:  /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
                12:  /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
                13:  /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
                14:  /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                              SUMMARIES
                %
Result          Query
  No.   Score   Match Length DB  ID                       Description
-----------------------------------------------------------------------------
    1    6157   100.0   3702  10  US-09-873-409-13         Sequence 13, Appl
    2  5988.5    97.2   3621  10  US-09-873-409-14         Sequence 14, Appl
    3    5328    86.5   3177  10  US-09-873-409-12         Sequence 12, Appl
    4  4561.5    74.1   2856  10  US-09-873-409-10         Sequence 10, Appl
    5  3530.5    57.3   3860  10  US-09-866-866A-1         Sequence 1, Appli
    6  3525.5    57.2   4317   9  US-10-044-671-1          Sequence 1, Appli
    7  3521.5    57.2   3860  10  US-09-866-866A-3         Sequence 3, Appli
    8  3521.5    57.2   8630  10  US-09-306-417-1          Sequence 1, Appli
    9  3521.5    57.2   8630  10  US-09-306-417-2          Sequence 2, Appli
   10    3517    57.1   4425  10  US-09-769-097-3          Sequence 3, Appli
   11    3515    57.1   4369  10  US-09-769-097-1          Sequence 1, Appli
   12    3514    57.1   4643   9  US-10-072-621-2          Sequence 2, Appli
   13    3503    56.9   4189  10  US-09-866-866A-5         Sequence 5, Appli
   14    3499    56.8   4788  10  US-09-866-866A-7         Sequence 7, Appli
   15    3481    56.5   3924  10  US-09-880-107-2299       Sequence 2299, Ap
   16    3456    56.1   2066  10  US-09-873-409-9          Sequence 9, Appli
   17    3453    56.1   4254  10  US-09-917-800A-1424      Sequence 1424, Ap
   18    3405    55.3   3912  10  US-09-917-800A-1560      Sequence 1560, Ap
   19  2811.5    45.6   5010  10  US-09-917-800A-483       Sequence 483, App
   20    2697    43.8   2021  10  US-09-873-409-15         Sequence 15, Appl
   21  2528.5    41.1   1940  10  US-09-873-409-16         Sequence 16, Appl
   22    2358    38.3   4175  10  US-09-749-340-3          Sequence 3, Appli
   23    2328    37.8   3861   9  US-09-938-842A-2237      Sequence 2237, Ap
   24    2259    36.7   4002   9  US-09-758-828-1          Sequence 1, Appli
   25    2259    36.7   4002   9  US-09-758-828-3          Sequence 3, Appli
   26  2129.5    34.6   3861   9  US-09-938-842A-263       Sequence 263, App
   27  2039.5    33.1   3512  10  US-09-749-340-7          Sequence 7, Appli
   28    1987    32.3   4653   9  US-10-101-388-2          Sequence 2, Appli
   29    1636    26.6   3792   9  US-09-882-694-10         Sequence 10, Appl
   30    1617    26.3   2698  10  US-09-749-340-5          Sequence 5, Appli
   31  1538.5    25.0   2681  10  US-09-749-340-8          Sequence 8, Appli
   32  1526.5    24.8   8036   9  US-10-101-388-1          Sequence 1, Appli
   33  1501.5    24.4   3999   9  US-09-882-694-9          Sequence 9, Appli
   34  1101.5    17.9   1175  10  US-09-873-409-11         Sequence 11, Appl
   35    1080    17.5   1810  10  US-09-749-340-4          Sequence 4, Appli
   36  1066.5    17.3   7296  10  US-09-070-927A-59        Sequence 59, Appl
   37    1027    16.7   3741   9  US-09-738-626-1251       Sequence 1251, Ap
c  38  1008.5    16.4   9047  10  US-09-070-927A-12        Sequence 12, Appl
   39   981.5    15.9  12438  10  US-09-070-927A-173       Sequence 173, App
   40   930.5    15.1 640681  10  US-09-790-988-1          Sequence 1, Appli
c  41   902.5    14.7   3601  10  US-09-070-927A-255       Sequence 255, App
c  42   901.5    14.6   6415  10  US-09-070-927A-275       Sequence 275, App
   43   770.5    12.5   2298   9  US-10-156-239-18         Sequence 18, Appl
   44   770.5    12.5   2298  10  US-09-795-693-18         Sequence 18, Appl
   45   770.5    12.5   3408   9  US-10-156-239-16         Sequence 16, Appl

                              ALIGNMENTS

RESULT 1
US-09-873-409-13
; Sequence 13, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
;   LENGTH: 3702
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: No. US20020037522A1e
;   LOCATION: (723)..(723)
;   OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-13

Alignment Scores:
Pred. No.:              0              Length:         3702
Score:                  6157.00        Matches:        1222
Percent Similarity:     100.00%        Conservative:   0
```

```
Best Local Similarity:  100.00%        Mismatches:   0
Query Match:            99.97%         Indels:       0
DB:                     10             Gaps:         0

US-09-873-409-5 (1-1222) x US-09-873-409-13 (1-3702)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     34 ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 93

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153

Qy     41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    154 TCTTTCTTCAGGTTGACCCTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGT 213

Qy     61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    214 TACATACAGATTTCCTTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAA 273

Qy     81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    274 CAGTTTTTTCATTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333

Qy    101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    334 GAACTTAACACTCGCATGACAGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATT 393

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAG 453

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    454 GGCTGGAAACTCACCCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA 513

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAG 633

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    634 GAGAAAGAACTTCAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTTGGTTTTATTTT 693

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    694 CCCCAGTGGCTACTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGAATCTCAAAGAT 753

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    754 GCAAAGGATTTTGGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTAC 813

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    814 TTCTTTATGAATGGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAAT 873

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    874 GGAGAACCTGGATATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGC 933

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    934 AGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCT 993

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    994 GCCTTTCATATTTTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCT 1053

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1054 GGATATAAACCTGAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTAT 1113

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1114 CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAG 1173

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1174 ACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG 1233

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1234 AGGTTATATGATCCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1293

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1294 AATGTGCGGCATTATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGG 1353

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1354 ACCACCATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1413

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1414 AGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAAT 1473

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1474 ACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATT 1533

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1534 GCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTG 1593

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1594 GATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTGGAGAAGGCGAGCAAAGGTCGGACT 1653

Qy    541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1654 ACAATCGTGGTAGCACACCGACTTTCTACTATTCGAAGTGCAGATTTGATTGTGACCCTA 1713

Qy    561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1714 AAGGATGGAATGCTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTA 1773

Qy    581 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1774 TATTATTCACTTGTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATG 1833

Qy    601 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1834 ACATATTCTACTGAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAG 1893

Qy    621 SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGlu 640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1894 TCAGACTTCATTGACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAA 1953

Qy    641 ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGly 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1954 GTCTCTCTATTAAAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGG 2013

Qy    661 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2014 ACATTGGCTTCTGTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAA 2073

Qy    681 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2074 ATTATAACCATGTTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTAT 2133
```

```
Qy      701 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2134 TCCATGATATTCGTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTA 2193

Qy      721 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 740
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2194 TTTTACGGCAGAGCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCC 2253

Qy      741 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 760
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2254 ATGTTATATCAGGATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACA 2313

Qy      761 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2314 ACAATATTAGCCATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTC 2373

Qy      781 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrp 800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2374 TTAACACAAAATGCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGG 2433

Qy      801 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2434 GAGATGACATTCCTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAA 2493

Qy      821 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2494 ACCGCAGCAATGACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAG 2553

Qy      841 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2554 ATAGCAACTGAAGCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCC 2613

Qy      861 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2614 TTCGAGCAAATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCA 2673

Qy      881 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2674 CAGATTATTGGAAGCTGTTATGCATTCAGCCATGCCTTTATATATTTGCCTATGCAGCA 2733

Qy      901 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2734 GGGTTTCGATTTGGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTC 2793

Qy      921 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2794 ATAGTTTTTACTGCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCT 2853

Qy      941 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2854 CCTGAATATTCCAAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAA 2913

Qy      961 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2914 CCAAATATAGACAGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTA 2973

Qy      981 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 1000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2974 GAGTTTCGAGAAGTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGC 3033

Qy     1001 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3034 TTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGG 3093

Qy     1021 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 1040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3094 AAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTG 3153

Qy     1041 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 1060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3154 TTTGATGGTGTGGATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATC 3213

Qy     1061 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3214 GTTCCTCAAGAGCCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGAC 3273

Qy     1081 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 1100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3274 AACAGCCGTGTGGTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCAT 3333

Qy     1101 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 1120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3334 TCTTTTATTGAAGGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAG 3393

Qy     1121 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3394 CTTTCTGGCGGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAA 3453

Qy     1141 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 1160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3454 ATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAG 3513

Qy     1161 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 1180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3514 CATGCCCTTGATAAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCT 3573

Qy     1181 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3574 GCAATTCAGAACGCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGA 3633

Qy     1201 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3634 ACTCATCAAGAGCTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCA 3693

Qy     1221 ValGln 1222
            ||||||
Db     3694 GTGCAG 3699
```

RESULT 2
US-09-873-409-14
; Sequence 14, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
; LENGTH: 3621
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-873-409-14

Alignment Scores:

Pred. No.:	0	Length:	3621
Score:	5988.50	Matches:	1195
Percent Similarity:	97.79%	Conservative:	0
Best Local Similarity:	97.79%	Mismatches:	0
Query Match:	97.23%	Indels:	27
DB:	10	Gaps:	1

US-09-873-409-5 (1-1222) x US-09-873-409-14 (1-3621)

```
Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       34 ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 93

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153
```

```
Qy      41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     154 TCTTTCTTCAGGTTGACCCTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGT 213

Qy      61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     214 TACATACAGATTTCCTTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAA 273

Qy      81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     274 CAGTTTTTTCATTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333

Qy     101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     334 GAACTTAACACTCGCATGACAGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATT 393

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAG 453

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     454 GGCTGGAAACTCACCCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA 513

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAG 633

Qy     201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
           ||||||||||||||
Db     634 GAGAAAGAACTTCAA--------------------------------------------- 648

Qy     221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                               ||||||||||||||||||||||||
Db     649 ------------------------------------AGGTATACACAGAATCTCAAAGAT 672

Qy     241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     673 GCAAAGGATTTTGGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTAC 732

Qy     261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     733 TTCTTTATGAATGGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAAT 792

Qy     281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     793 GGAGAACCTGGATATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGC 852

Qy     301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     853 AGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCT 912

Qy     321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     913 GCCTTTCATATTTTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCT 972

Qy     341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     973 GGATATAAACCTGAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTAT 1032

Qy     361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1033 CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAG 1092

Qy     381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1093 ACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG 1152

Qy     401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1153 AGGTTATATGATCCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1212

Qy     421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1213 AATGTGCGGCATTATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGG 1272

Qy     441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1273 ACCACCATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1332

Qy     461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1333 AGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAAT 1392

Qy     481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1393 ACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATT 1452

Qy     501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1453 GCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTG 1512

Qy     521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1513 GATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTGGAGAAGGCGAGCAAAGGTCGGACT 1572

Qy     541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1573 ACAATCGTGGTAGCACACCGACTTTCTACTATTCGAAGTGCAGATTTGATTGTGACCCTA 1632

Qy     561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1633 AAGGATGGAATGCTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTA 1692

Qy     581 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1693 TATTATTCACTTGTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATG 1752

Qy     601 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1753 ACATATTCTACTGAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAG 1812

Qy     621 SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGlu 640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1813 TCAGACTTCATTGACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAA 1872

Qy     641 ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGly 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1873 GTCTCTCTATTAAAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGG 1932

Qy     661 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1933 ACATTGGCTTCTGTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAA 1992

Qy     681 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1993 ATTATAACCATGTTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTAT 2052

Qy     701 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2053 TCCATGATATTCGTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTA 2112

Qy     721 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2113 TTTTACGGCAGAGCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCC 2172

Qy     741 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2173 ATGTTATATCAGGATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACA 2232

Qy     761 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     2233 ACAATATTAGCCATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTC 2292

Qy      781 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrp 800
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2293 TTAACACAAAATGCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGG 2352

Qy      801 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 820
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2353 GAGATGACATTCCTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAA 2412

Qy      821 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2413 ACCGCAGCAATGACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAG 2472

Qy      841 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 860
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2473 ATAGCAACTGAAGCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCC 2532

Qy      861 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 880
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2533 TTCGAGCAAATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCA 2592

Qy      881 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2593 CAGATTATTGGAAGCTGTTATGCATTCAGCCATGCCTTTATATATTTTGCCTATGCAGCA 2652

Qy      901 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 920
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2653 GGGTTTCGATTTGGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTC 2712

Qy      921 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 940
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2713 ATAGTTTTTACTGCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCT 2772

Qy      941 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2773 CCTGAATATTCCAAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAA 2832

Qy      961 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 980
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2833 CCAAATATAGACAGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTA 2892

Qy      981 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 1000
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2893 GAGTTTCGAGAAGTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGC 2952

Qy     1001 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2953 TTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGG 3012

Qy     1021 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 1040
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3013 AAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTG 3072

Qy     1041 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 1060
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3073 TTTGATGGTGTGGATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATC 3132

Qy     1061 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3133 GTTCCTCAAGAGCCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGAC 3192

Qy     1081 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 1100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3193 AACAGCCGTGTGGTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCAT 3252

Qy     1101 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 1120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3253 TCTTTTATTGAAGGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAG 3312

Qy     1121 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3313 CTTTCTGGCGGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAA 3372

Qy     1141 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 1160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3373 ATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAG 3432

Qy     1161 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 1180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3433 CATGCCCTTGATAAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCT 3492

Qy     1181 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3493 GCAATTCAGAACGCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGA 3552

Qy     1201 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3553 ACTCATCAAGAGCTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCA 3612

Qy     1221 ValGln 1222
            ||||||
Db     3613 GTGCAG 3618

RESULT 3
US-09-873-409-12
; Sequence 12, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
;   LENGTH: 3177
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: No. US20020037522A1e
;   LOCATION: (198)..(198)
;   OTHER INFORMATION: n at position 198 represents any nucleotide (A, T, C or G)
US-09-873-409-12

Alignment Scores:
Pred. No.:              0               Length:         3177
Score:                  5328.00         Matches:        1058
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            86.51%          Indels:         0
DB:                     10              Gaps:           0

US-09-873-409-5 (1-1222) x US-09-873-409-12 (1-3177)

Qy      165 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 184
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 ATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCA 60

Qy      185 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 204
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 GAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAGGAGAAAGAACTT 120

Qy      205 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 224
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      121 CAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTTGGTTTTATTTCCCCAGTGGCTA 180

Qy      225 LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 244
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      181 CTAAGTTGTGTTCTGTTNTTGTAAGGTATACACAGAATCTCAAAGATGCAAAGGATTTT 240

Qy      245 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 264
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     241 GGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAAT 300

Qy     265 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 284
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGA 360

Qy     285 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 304
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATT 420

Qy     305 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 324
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATT 480

Qy     325 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 344
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCT 540

Qy     345 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 364
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCA 600

Qy     365 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 384
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 TCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTG 660

Qy     385 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 404
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 720

Qy     405 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 424
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 CCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCAT 780

Qy     425 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 444
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 TATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGT 840

Qy     445 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 464
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 AACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGG 900

Qy     465 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 484
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGG 960

Qy     485 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 504
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTA 1020

Qy     505 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 524
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGC 1080

Qy     525 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 544
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 AAGTCAGCTGTTCAAGCTGCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTA 1140

Qy     545 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 564
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 GCACACCGACTTTCTACTATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATG 1200

Qy     565 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 584
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 CTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTT 1260

Qy     585 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 604
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACT 1320

Qy     605 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 624
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 GAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATT 1380

Qy     625 AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 644
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 GACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTA 1440

Qy     645 LysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 664
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 AAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCT 1500

Qy     665 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 684
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 GTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATG 1560

Qy     685 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 704
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 TTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTC 1620

Qy     705 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 724
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 GTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGA 1680

Qy     725 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 744
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 GCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAG 1740

Qy     745 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 764
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 GATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCC 1800

Qy     765 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 784
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801 ATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAAT 1860

Qy     785 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 804
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861 GCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTC 1920

Qy     805 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 824
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921 CTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATG 1980

Qy     825 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 844
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1981 ACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAA 2040

Qy     845 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 864
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2041 GCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATG 2100

Qy     865 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 884
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2101 TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGA 2160

Qy     885 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 904
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2161 AGCTGTTATGCATTCAGCCATGCCTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTT 2220

Qy     905 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 924
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2221 GGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACT 2280

Qy     925 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 944
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2281 GCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCC 2340

Qy     945 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 964
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2341 AAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGAC 2400

Qy     965 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 984
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2401 AGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAA 2460
```

```
Qy     985 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 1004
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2461 GTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGT 2520

Qy    1005 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 1024
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2521 ATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCT 2580

Qy    1025 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 1044
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2581 GTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTG 2640

Qy    1045 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 1064
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2641 GATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAG 2700

Qy    1065 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 1084
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2701 CCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTG 2760

Qy    1085 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 1104
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2761 GTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAA 2820

Qy    1105 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 1124
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2821 GGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGC 2880

Qy    1125 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 1144
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2881 CAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTG 2940

Qy    1145 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1164
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2941 GATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGAT 3000

Qy    1165 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1184
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3001 AAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAC 3060

Qy    1185 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1204
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3061 GCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAG 3120

Qy    1205 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1222
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3121 CTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3174
```

```
RESULT 4
US-09-873-409-10
; Sequence 10, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
;   LENGTH: 2856
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-409-10

Alignment Scores:
Pred. No.:               0            Length:         2856
Score:                   4561.50      Matches:        912
Percent Similarity:      97.96%       Conservative:   0
Best Local Similarity:   97.96%       Mismatches:     1
Query Match:             74.06%       Indels:         18
DB:                      10           Gaps:           1

US-09-873-409-5 (1-1222) x US-09-873-409-10 (1-2856)

Qy     293 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHis-PheGl 312
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     114 GTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATTGGAGCAGCAGTCCCTCATTATT-- 171

Qy     312 uThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPr 332
                                                             ||||||||||
Db     172 --------------------------------------------------GATAAGAAACC 182

Qy     332 oSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGl 352
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 CAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGA 242

Qy     352 uPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLe 372
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 ATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCT 302

Qy     372 uAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLy 392
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 GAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAA 362

Qy     392 sSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVa 412
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     363 GAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGT 422

Qy     412 lAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVa 432
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     423 GGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTGGT 482

Qy     432 lSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAs 452
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     483 TAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGAGA 542

Qy     452 pAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIl 472
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     543 TGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTTAT 602

Qy     472 eMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGl 492
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     603 CATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGG 662

Qy     492 yGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIl 512
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     663 AGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGAT 722

Qy     512 eLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLe 532
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     723 TTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACT 782

Qy     532 uGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleAr 552
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     783 GGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTAGCACACCGACTTTCTACTATTCG 842

Qy     552 gSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAl 572
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     843 AAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATGCTGGCGGAGAAAGGAGCACATGC 902

Qy     572 aGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLy 592
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     903 TGAACTAATGGCAAAACGAGGTCTATATTATTCACTTGTGATGTCACAGGATATTAAAAA 962

Qy     592 sAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPr 612
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     963 AGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCACTTCC 1022

Qy     612 oLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGl 632
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1023 TCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCCACCCA 1082
```

```
Qy     632 nSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPr 652
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1083 ATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAACAAGCC 1142

Qy     652 oGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPr 672
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1143 TGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCTGTTCTAAATGGAACTGTTCATCC 1202

Qy     672 oValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrTh 692
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1203 AGTATTTTCCATCATCTTTGCAAAAATTATAACCATGTTTGGAAATAATGATAAAACCAC 1262

Qy     692 rLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPh 712
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1263 ATTAAAGCATGATGCAGAAATTTATTCCATGATATTCGTCATTTTGGGTGTTATTTGCTT 1322

Qy     712 eValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetAr 732
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1323 TGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGAGCAGGGGAAATTTTAACGATGAG 1382

Qy     732 gLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLy 752
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1383 ATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAGGATATTGCCTGGTTTGATGAAAA 1442

Qy     752 sGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGl 772
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1443 GGAAAACAGCACAGGAGGCTTGACAACAATATTAGCCATAGATATAGCACAAATTCAAGG 1502

Qy     772 yAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVa 792
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1503 AGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAATGCAACTAACATGGGACTTTCAGT 1562

Qy     792 lIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVa 812
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1563 TATCATTTCCTTTATATATGGATGGGAGATGACATTCCTGATTCTGAGTATTGCTCCAGT 1622

Qy     812 lLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLy 832
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1623 ACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATGACTGGATTTGCCAACAAAGATAA 1682

Qy     832 sGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIl 852
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1683 GCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAGAATATACGTACTAT 1742

Qy     852 eValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGl 872
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1743 AGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATGTATGAAGAGATGCTTCAGACTCA 1802

Qy     872 nHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAl 892
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1803 ACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTATGCATTCAGCCATGC 1862

Qy     892 aPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGl 912
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1863 CTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTTGGAGCCTATTTAATTCAAGCTGG 1922

Qy     912 yArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAl 932
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1923 ACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACTGCAATTGCATATGGAGCTATGGC 1982

Qy     932 aIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHi 952
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1983 CATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCCAAAGCCAAATCGGGGGCTGCGCA 2042

Qy     952 sLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLy 972
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2043 TCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGACAGCCGCAGTCAAGAAGGGAAAAA 2102

Qy     972 sProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysAr 992
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2103 GCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTCTTCTATCCATGTCG 2162

Qy     992 gProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAl 1012
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2163 CCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGC 2222

Qy    1012 aPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTy 1032
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2223 ATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTA 2282

Qy    1032 rAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGl 1052
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2283 TGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTGGATGCAAAAGAATTGAATGTACA 2342

Qy    1052 nTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIl 1072
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2343 GTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAGCCTGTGCTCTTCAACTGCAGCAT 2402

Qy    1072 eAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGl 1092
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2403 TGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTGGTGCCATTAGATGAGATCAAAGA 2462

Qy    1092 uAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnTh 1112
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2463 AGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAAGGTCTCCCTGAGAAATACAACAC 2522

Qy    1112 rGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAl 1132
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2523 ACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCTATTGC 2582

Qy    1132 aArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAs 1152
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2583 AAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGA 2642

Qy    1152 pAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCy 1172
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2643 TAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGATAAAGCCAGGACGGGAAGGACATG 2702

Qy    1172 sLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHi 1192
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2703 CCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGGTTCTGCA 2762

Qy    1192 sAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTy 1212
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2763 CAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAGCTCCTGAGAAATCGAGACATATA 2822

Qy    1212 rPheLysLeuValAsnAlaGlnSerValGln 1222
           ||||||||||||||||||||||||||||||
Db    2823 TTTTAAGTTAGTGAATGCACAGTCAGTGCAG 2853
```

```
RESULT 5
US-09-866-866A-1
; Sequence 1, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
;  APPLICANT: Sorrentino, Brian
;  APPLICANT:  Schuetz, John
;  TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
;  FILE REFERENCE: 1340-1-021CIP2
;  CURRENT APPLICATION NUMBER: US/09/866,866A
;  CURRENT FILING DATE:  2001-08-30
;  PRIOR APPLICATION NUMBER: 09/584,586
;  PRIOR FILING DATE: 2000-05-31
;  PRIOR APPLICATION NUMBER: PCT/US99/11825
;  PRIOR FILING DATE: 1999-05-27
;  PRIOR APPLICATION NUMBER: 60/086,988
;  PRIOR FILING DATE: 1998-05-28
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 3860
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-866A-1

Alignment Scores:
```

```
Pred. No.:               4.44e-314   Length:        3860
Score:                   3530.50     Matches:       682
Percent Similarity:      74.00%      Conservative:  246
Best Local Similarity:   54.39%      Mismatches:    263
Query Match:             57.32%      Indels:        63
DB:                      10          Gaps:          10

US-09-873-409-5 (1-1222) x US-09-866-866A-1 (1-3860)

Qy     1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db   151 ATGGTGGTGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210

Qy    21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db   211 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 270

Qy    32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db   271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330

Qy    44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db   331 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 390

Qy    62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db   391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450

Qy    82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db   451 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 510

Qy   102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db   511 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGGTATTGGTGACAAAATT 570

Qy   121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db   571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 630

Qy   141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db   631 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

Qy   161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db   691 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy   181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db   751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810

Qy   201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db   811 AAGAAAGAACTTGAA--------------------------------------------- 825

Qy   221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db   826 -----------------------------------AGGTACAACAAAAATTTAGAAGAA 849

Qy   241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db   850 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 909

Qy   261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db   910 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 969

Qy   281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db   970 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1023

Qy   301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db  1024 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1083

Qy   321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db  1084 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1143

Qy   341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db  1144 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1203

Qy   361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db  1204 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1263

Qy   381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db  1264 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1323

Qy   401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db  1324 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1383

Qy   421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db  1384 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1443

Qy   441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db  1444 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1503

Qy   461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db  1504 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1563

Qy   481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db  1564 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 1623

Qy   501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db  1624 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 1683

Qy   521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db  1684 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACC 1743

Qy   541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db  1744 ACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC 1803

Qy   561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db  1804 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATT 1863

Qy   581 TyrTyrSerLeuValMetSerGln------------------AspIleLysLysAlaAsp 594
Db  1864 TACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 1923

Qy   595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
Db  1924 GATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGT 1983

Qy   615 SerVal------------------LysSerIleLys---SerAspPheIleAspLysAla 627
Db  1984 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAG 2043

Qy   628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
Db  2044 CTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATG 2103

Qy   648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
```

```
Db    2104 AAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAAT 2163

Qy     668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
Db    2164 GGAGGCCTGCAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGA 2223

Qy     688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db    2224 ATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCC 2283

Qy     707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db    2284 CTTGGAATTATTTCTTTTATTACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGA 2343

Qy     727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db    2344 GAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTG 2403

Qy     747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
Db    2404 AGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGAT 2463

Qy     767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db    2464 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCA 2523

Qy     787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
Db    2524 AATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTC 2583

Qy     807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db    2584 TTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGA 2643

Qy     827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db    2644 CAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 2703

Qy     847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db    2704 GAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCT 2763

Qy     867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db    2764 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 2823

Qy     887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db    2824 TTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCC 2883

Qy     907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db    2884 TACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTT 2943

Qy     927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db    2944 GTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCC 3003

Qy     947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db    3004 AAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTAC 3063

Qy     967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db    3064 AGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTA 3123

Qy     987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db    3124 TTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAG 3183

Qy    1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
Db    3184 AAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAG 3243

Qy    1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
Db    3244 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATA 3303

Qy    1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
Db    3304 AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATC 3363

Qy    1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db    3364 CTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCA 3423

Qy    1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db    3424 CAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTG 3483

Qy    1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
Db    3484 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAA 3543

Qy    1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
Db    3544 CAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAA 3603

Qy    1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
Db    3604 GCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCC 3663

Qy    1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db    3664 AGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGAC 3723

Qy    1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
Db    3724 TTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTG 3783

Qy    1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db    3784 GCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3825
```

RESULT 6
US-10-044-671-1
; Sequence 1, Application US/10044671
; Patent No. US20020177147A1
; GENERAL INFORMATION:
; APPLICANT: Washington State University Research Foundation
; APPLICANT: Mealey, Katrina
; APPLICANT: Bentjen, Steven
; TITLE OF INVENTION: MDR1 VARIANTS AND METHODS FOR THEIR USE
; FILE REFERENCE: 4630-61733
; CURRENT APPLICATION NUMBER: US/10/044,671
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/261,578
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/314,829
; PRIOR FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4317
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (70)..(3912)
; OTHER INFORMATION:
US-10-044-671-1

```
Alignment Scores:
Pred. No.:          1.51e-313     Length:      4317
Score:              3525.50       Matches:     680
```

```
Percent Similarity:       73.05%        Conservative:   239
Best Local Similarity:    54.05%        Mismatches:     270
Query Match:              57.24%        Indels:         69
DB:                       9             Gaps:           9

US-09-873-409-5 (1-1222) x US-10-044-671-1 (1-4317)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db      217 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 276

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
Db      277 GTTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 336

Qy       35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
Db      337 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 396

Qy       45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db      397 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 456

Qy       60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db      457 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 516

Qy       80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db      517 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 576

Qy      100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db      577 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 636

Qy      119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db      637 AAAGTTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 696

Qy      139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db      697 ACACCTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 756

Qy      159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db      757 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 816

Qy      179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db      817 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 876

Qy      199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
Db      877 GGACAAAAGAAAGAACTTGAA--------------------------------------- 897

Qy      219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
Db      898 ------------------------------------------AGGTACAACAAAAATTTA 915

Qy      239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
Db      916 GAAGAAGCTAAAAGAATTGGGATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCC 975

Qy      259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
Db      976 GCTTTCTTATTGATCTATGCATCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTC 1035

Qy      279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
Db     1036 CTCTCCAGTGAA------TATACTATTGGACAGGTACTCACTGTCTTCTTTTCTGTATTA 1089

Qy      299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
Db     1090 ATTGGGGCTTTTAGTATTGGACAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGA 1149

Qy      319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
Db     1150 GGAGCAGCTTATGAAATCTTCAAGATAATTGACAATAAACCAAGCATTGACAGCTATTCG 1209

Qy      339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
Db     1210 AAGAGTGGACATAAACCAGATAATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTC 1269

Qy      359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
Db     1270 AGTTACCCTTCTCGAAAAGAAGTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGT 1329

Qy      379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
Db     1330 GGGCAGACAGTGGCGCTGGTTGGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTG 1389

Qy      399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
Db     1390 ATGCAGAGGCTCTATGACCCCACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGG 1449

Qy      419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db     1450 ACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTG 1509

Qy      439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
Db     1510 TTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAG 1569

Qy      459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
Db     1570 ATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTTATCATGAAACTACCTAATAAA 1629

Qy      479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db     1630 TTTGACACTCTGGTTGGAGAGAGAGGGGCCCGGCTGAGTGGTGGACAGAAACAGAGAATC 1689

Qy      499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db     1690 GCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCA 1749

Qy      519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 538
Db     1750 GCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGC 1809

Qy      539 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558
Db     1810 CGGACTACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCCGATGTCATTGCT 1869

Qy      559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578
Db     1870 GGTTTTGATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAG 1929

Qy      579 GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 598
Db     1930 GGCATTTACTTCAAACTTGTCACAATGCAG---ACAAGAGGAAATGAAATTGAGTTAGAA 1986

Qy      599 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 618
Db     1987 AATGCCACTGGTGAATCCAAAAGTGAAAGTGATGCCTTGGAAATG---TCTCCAAAAGAT 2043

Qy      619 IleLysSerAspPheIle------------------------------------------ 624
Db     2044 TCAGGGTCCAGTTTAATAAAAAGAAGATCAACTCGCAGGAGTATACATGCACCACAAGGC 2103

Qy      625 ---AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643
Db     2104 CAAGACAGAAAGCTTGGTACAAAAGAGGACTTGAATGAGAATGTTCCTTCAGTTTCCTTC 2163

Qy      644 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 663
Db     2164 TGGAGGATTCTGAAGCTGAACTCAACTGAATGGCCTTATTTTGTGGTTGGTATATTTTGT 2223
```

```
Qy    664 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 683
Db   2224 GCTATTATAAACGGAGGCCTGCAACCAGCATTTTCAATAATATTTTCAAGGATTATAGGG 2283

Qy    684 MetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 702
Db   2284 ATCTTTACCCGAGATGAGGATCCTGAAACAAAACGACAGAATAGTAACATGTTTTCTGTA 2343

Qy    703 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722
Db   2344 TTGTTTCTAGTCCTTGGAATTATTTCTTTTATTACATTTTTCCTCCAGGGCTTCACATTT 2403

Qy    723 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 742
Db   2404 GGCAAAGCTGGGGAGATCCTCACTAAGCGGCTTCGATACATGGTTTTCAGATCCATGCTG 2463

Qy    743 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIle 762
Db   2464 AGACAGGATGTCAGCTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACAACCAGG 2523

Qy    763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 782
Db   2524 CTTGCCAATGATGCGGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTCATTACC 2583

Qy    783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 802
Db   2584 CAGAATATAGCAAATCTTGGGACAGGCATTATTATATCCTTAATCTATGGTTGGCAATTA 2643

Qy    803 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 822
Db   2644 ACACTTTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAA 2703

Qy    823 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 842
Db   2704 ATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCT 2763

Qy    843 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 862
Db   2764 ACAGAAGCCATCGAAAACTTCCGAACTGTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAA 2823

Qy    863 GlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 882
Db   2824 TACATGTATGCACAGAGTTTGCAAGTACCATACAGAAACTCTTTGAGGAAAGCACACATC 2883

Qy    883 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 902
Db   2884 TTCGGGGTCTCATTTTCTATCACCCAGGCAATGATGTATTTTTCCTATGCTGGCTGTTTC 2943

Qy    903 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 922
Db   2944 CGGTTTGGTGCCTACTTGGTGGCAAATGAGTTCATGAACTTTCAGGATGTTCTTTTGGTA 3003

Qy    923 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 942
Db   3004 TTCTCAGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAGGTCAGTTCATTTGCTCCTGAC 3063

Qy    943 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 962
Db   3064 TATGCCAAAGCCAAAGTATCAGCAGCCCACGTCATCATGATCATTGAAAAAAGCCCTCTG 3123

Qy    963 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 982
Db   3124 ATTGACAGCTACAGCCCTCACGGCCTCAAGCCAAATACGTTGGAAGGAAATGTGACATTT 3183

Qy    983 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002
Db   3184 AATGAGGTCGTGTTCAACTATCCCACTCGACCAGACATCCCCGTGCTCCAGGGGCTGAGC 3243

Qy   1003 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 1022
Db   3244 CTCGAGGTGAAGAAGGGCCAGACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGC 3303

Qy   1023 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1042
Db   3304 ACAGTTGTTCAGCTCCTAGAGCGCTTCTATGACCCCTTGGCTGGTTCAGTGCTAATTGAT 3363

Qy   1043 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 1062
Db   3364 GGCAAAGAGATAAAGCACCTGAATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCT 3423

Qy   1063 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082
Db   3424 CAGGAGCCCATCCTGTTTGACTGCAGCATTGCCGAGAACATTGCCTATGGAGACAACAGC 3483

Qy   1083 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1102
Db   3484 CGGGTCGTATCACATGAAGAGATTATGCAGGCAGCCAAGGAGGCCAACATACACCACTTC 3543

Qy   1103 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122
Db   3544 ATCGAGACACTCCCTGAGAAATACAACACCAGAGTAGGAGACAAAGGAACCCAGCTCTCT 3603

Qy   1123 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 1142
Db   3604 GGTGGCCAGAAACAGCGCATTGCCATAGCTCGCGCTCTTGTTAGACAGCCTCATATTTTG 3663

Qy   1143 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 1162
Db   3664 CTTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCC 3723

Qy   1163 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 1182
Db   3724 CTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATCGCCCACCGCTTGTCCACCATC 3783

Qy   1183 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 1202
Db   3784 CAGAATGCAGATTTAATAGTGGTGTTTCAGAATGGCAAAGTCAAGGAGCATGGCACACAT 3843

Qy   1203 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db   3844 CAACAGCTGCTGGCTCAGAAAGGCATCTATTTTTCCATGATCAGTGTCCAGGCT 3897
```

```
RESULT 7
US-09-866-866A-3
; Sequence 3, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
;  APPLICANT: Sorrentino, Brian
;  APPLICANT:  Schuetz, John
;  TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
;  FILE REFERENCE: 1340-1-021CIP2
;  CURRENT APPLICATION NUMBER: US/09/866,866A
;  CURRENT FILING DATE:  2001-08-30
;  PRIOR APPLICATION NUMBER: 09/584,586
;  PRIOR FILING DATE: 2000-05-31
;  PRIOR APPLICATION NUMBER: PCT/US99/11825
;  PRIOR FILING DATE: 1999-05-27
;  PRIOR APPLICATION NUMBER: 60/086,988
;  PRIOR FILING DATE: 1998-05-28
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
;   LENGTH: 3860
;   TYPE: DNA
;   ORGANISM: homo sapiens
US-09-866-866A-3

Alignment Scores:
Pred. No.:              2.98e-313     Length:        3860
Score:                  3521.50       Matches:       681
Percent Similarity:     73.92%        Conservative:  246
Best Local Similarity:  54.31%        Mismatches:    264
Query Match:            57.18%        Indels:        63
DB:                     10            Gaps:          10
```

```
US-09-873-409-5 (1-1222) x US-09-866-866A-3 (1-3860)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    151 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210

Qy     21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db    211 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 270

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe----------------- 43
Db    271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    331 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 390

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    451 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 510

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    511 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 570

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 630

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    631 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    691 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db    811 AAGAAAGAACTTGAA--------------------------------------------- 825

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db    826 ------------------------------------AGGTACAACAAAAATTTAGAAGAA 849

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db    850 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 909

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db    910 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 969

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db    970 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1023

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db   1024 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1083

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   1084 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1143

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   1144 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1203

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   1204 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1263

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   1264 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1323

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   1324 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1383

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   1384 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1443

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   1444 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1503

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   1504 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1563

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   1564 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 1623

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   1624 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 1683

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db   1684 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACC 1743

Qy    541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db   1744 ACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC 1803

Qy    561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db   1804 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATT 1863

Qy    581 TyrTyrSerLeuValMetSerGln-----------------AspIleLysLysAlaAsp 594
Db   1864 TACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 1923

Qy    595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
Db   1924 GATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGT 1983

Qy    615 SerVal-----------------LysSerIleLys---SerAspPheIleAspLysAla 627
Db   1984 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAG 2043

Qy    628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
Db   2044 CTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATG 2103

Qy    648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
Db   2104 AAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAAT 2163

Qy    668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
Db   2164 GGAGGCCTGCAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGA 2223
```

```
Qy    688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
             :::|||      |||   :::   ::::::  ::::::|||::::::|||:::
Db   2224 ATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCC 2283

Qy    707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
          ||||||:::|||   |||::::::::|||:::||||||        :::|||:::||||||
Db   2284 CTTGGAATTATTTCTTTTATTACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGA 2343

Qy    727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
          ||||||||||||   |||||||||::::::   |||::::::||||||   ||||||:::
Db   2344 GAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTG 2403

Qy    747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
          :::|||||||||:::   :::|||:::||||||   |||||||||   ||||||   |||
Db   2404 AGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGAT 2463

Qy    767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
                ||||||::::::||||||   |||||||||:::   |||:::|||||||||
Db   2464 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCA 2523

Qy    787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
          |||:::|||      :::||||||||||||||||||||||||||||::::::|||   |||:::
Db   2524 AATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTC 2583

Qy    807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
          |||:::|||   |||::::::|||:::   |||::::::|||         ::::::|||
Db   2584 TTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGA 2643

Qy    827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
            .|||   |||||||||:::||||||:::   ||||||||||||||||||||||||:::
Db   2644 CAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 2703

Qy    847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
          ||||||   ||||||:::||||||||||||:::|||:::   ||||||   ||||||
Db   2704 GAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCT 2763

Qy    867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
          :::   ||||||      :::|||||||||   :::||||||   |||   |||
Db   2764 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 2823

Qy    887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
          ::::::|||:::   |||   :::|||||||||:::||||||      |||||||||||||||
Db   2824 TTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCC 2883

Qy    907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
          ||||||:::               |||:::   |||   :::   :::||||||:::|||:::
Db   2884 TACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTT 2943

Qy    927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
             :::||||||||||||:::|||:::                  ||||||:::|||:::||||||
Db   2944 GTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCC 3003

Qy    947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
          |||      |||||||||:::         ::::::||||||   |||   |||||||||
Db   3004 AAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTAC 3063

Qy    967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
          |||   ||||||      |||:::|||   ||||||||||:::      |||   ||||||
Db   3064 AGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTA 3123

Qy    987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
          |||   ||||||   |||||||||:::   :::|||:::||||||||||||||      ::::::
Db   3124 TTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAG 3183

Qy   1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
          :::|||:::|||:::|||   |||||||||||||||||||||||||||||   ||||||
Db   3184 AAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAG 3243

Qy   1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
          ||||||:::|||   |||||||||:::   |||:::||||||   ||||||     :::
Db   3244 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATA 3303

Qy   1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
          |||   |||||||||||||||||||:::   :::   ||||||   |||||||||:::
Db   3304 AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATC 3363

Qy   1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
          ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3364 CTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCA 3423

Qy   1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
             :::||||||      ||||||   ||||||||||||:::||||||||   |||
Db   3424 CAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTG 3483

Qy   1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
          |||   ||||||:::|||:::||||||   ||||||   ||||||||||||||||||
Db   3484 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAA 3543

Qy   1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
          ||||||:::||||||||||||||||||:::::::::|||   ||||||||||||||||
Db   3544 CAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAA 3603

Qy   1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
          ||||||||||||||||||      :::||||||||||||||||||   ||||||||||||
Db   3604 GCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCC 3663

Qy   1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
          |||   ||||||||||||:::|||:::      ||||||||||||   |||||||||||||
Db   3664 AGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGAC 3723

Qy   1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
          ||||||||||||      ||||||::::::||||||   ||||||||||||:::||||||
Db   3724 TTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTG 3783

Qy   1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
                :::   |||||||||   :::|||:::   |||:::
Db   3784 GCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3825
```

RESULT 8
US-09-306-417-1
; Sequence 1, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 8630
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: proviral
;   OTHER INFORMATION: plasmid DNA
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(160)
;   OTHER INFORMATION: plasmid backbone (pUC)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (161)..(677)
;   OTHER INFORMATION: 5'-LTR
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (532)..(1219)
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: (1220)..(5062)
;   OTHER INFORMATION: m4 mdr-1 cDNA
;   FEATURE:
;   NAME/KEY: misc_feature

```
;   LOCATION: (5215)..(5774)
;   OTHER INFORMATION: 3'-LTR
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5775)..(8630)
;   OTHER INFORMATION: plasmid backbone (pUC)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(8630)
;   OTHER INFORMATION: retroviral expression vector SFbeta71m4
US-09-306-417-1

Alignment Scores:
Pred. No.:              9.87e-313     Length:         8630
Score:                  3521.50       Matches:        681
Percent Similarity:     73.92%        Conservative:   246
Best Local Similarity:  54.31%        Mismatches:     264
Query Match:            57.18%        Indels:         63
DB:                     10            Gaps:           10

US-09-873-409-5 (1-1222) x US-09-306-417-1 (1-8630)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db   1370 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 1429

Qy     21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db   1430 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 1489

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db   1490 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 1549

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db   1550 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 1609

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db   1610 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 1669

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db   1670 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 1729

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db   1730 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 1789

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db   1790 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1849

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db   1850 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1909

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db   1910 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1969

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db   1970 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 2029

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db   2030 AAGAAAGAACTTGAA--------------------------------------------- 2044

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db   2045 -----------------------------------AGGTACAACAAAAATTTAGAAGAA 2068

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db   2069 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 2128

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db   2129 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 2188

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db   2189 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 2242

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db   2243 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 2302

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   2303 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 2362

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   2363 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 2422

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   2423 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 2482

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   2483 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 2542

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   2543 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 2602

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   2603 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 2662

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   2663 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 2722

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   2723 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 2782

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   2783 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 2842

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   2843 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 2902

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db   2903 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACC 2962

Qy    541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db   2963 ACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC 3022

Qy    561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db   3023 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATT 3082

Qy    581 TyrTyrSerLeuValMetSerGln------------------AspIleLysLysAlaAsp 594
Db   3083 TACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 3142
```

```
Qy    595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
Db   3143 GATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGT 3202

Qy    615 SerVal------------------LysSerIleLys---SerAspPheIleAspLysAla 627
Db   3203 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAG 3262

Qy    628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
Db   3263 CTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATG 3322

Qy    648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
Db   3323 AAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAAT 3382

Qy    668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
Db   3383 GGAGGCCTGCAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGA 3442

Qy    688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db   3443 ATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCC 3502

Qy    707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db   3503 CTTGGAATTATTTCTTTTATTACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGA 3562

Qy    727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db   3563 GAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTG 3622

Qy    747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
Db   3623 AGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGAT 3682

Qy    767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db   3683 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCA 3742

Qy    787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
Db   3743 AATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTC 3802

Qy    807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db   3803 TTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGA 3862

Qy    827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db   3863 CAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 3922

Qy    847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db   3923 GAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCT 3982

Qy    867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db   3983 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 4042

Qy    887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db   4043 TTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCC 4102

Qy    907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db   4103 TACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTT 4162

Qy    927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db   4163 GTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCC 4222

Qy    947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db   4223 AAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTAC 4282

Qy    967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db   4283 AGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTA 4342

Qy    987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db   4343 TTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAG 4402

Qy   1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
Db   4403 AAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAG 4462

Qy   1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
Db   4463 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATT 4522

Qy   1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
Db   4523 AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATC 4582

Qy   1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db   4583 CTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCA 4642

Qy   1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db   4643 CAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTG 4702

Qy   1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
Db   4703 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAA 4762

Qy   1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
Db   4763 CAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAA 4822

Qy   1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
Db   4823 GCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCC 4882

Qy   1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db   4883 AGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGAC 4942

Qy   1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
Db   4943 TTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTG 5002

Qy   1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db   5003 GCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 5044
```

RESULT 9
US-09-306-417-2
; Sequence 2, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence

```
;    FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence: proviral
;    OTHER INFORMATION: plasmid DNA
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (1)..(8630)
;    OTHER INFORMATION: retroviral expression vector SFbeta91mSA1
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (1)..(160)
;    OTHER INFORMATION: plasmid backbone (pUC)
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (161)..(677)
;    OTHER INFORMATION: 5'-LTR
;    FEATURE:
;    NAME/KEY: 5'UTR
;    LOCATION: (532)..(1219)
;    FEATURE:
;    NAME/KEY: mat_peptide
;    LOCATION: (1220)..(5062)
;    OTHER INFORMATION: mSA1 mdr1 cDNA
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (5215)..(5774)
;    OTHER INFORMATION: 3'-LTR
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (5775)..(8630)
;    OTHER INFORMATION: plasmid backbone (pUC)
US-09-306-417-2

Alignment Scores:
Pred. No.:              9.87e-313    Length:         8630
Score:                  3521.50      Matches:        681
Percent Similarity:     73.92%       Conservative:   246
Best Local Similarity:  54.31%       Mismatches:     264
Query Match:            57.18%       Indels:         63
DB:                     10           Gaps:           10

US-09-873-409-5 (1-1222) x US-09-306-417-2 (1-8630)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db   1370 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 1429

Qy     21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db   1430 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 1489

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db   1490 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 1549

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db   1550 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 1609

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db   1610 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 1669

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db   1670 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 1729

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db   1730 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 1789

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db   1790 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1849

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db   1850 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1909

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db   1910 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1969

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db   1970 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 2029

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db   2030 AAGAAAGAACTTGAA--------------------------------------------- 2044

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db   2045 -----------------------------------AGGTACAACAAAAATTTAGAAGAA 2068

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db   2069 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 2128

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db   2129 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 2188

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db   2189 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 2242

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db   2243 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 2302

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   2303 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 2362

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   2363 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 2422

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   2423 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 2482

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   2483 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 2542

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   2543 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 2602

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   2603 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 2662

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   2663 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 2722

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   2723 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 2782

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   2783 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 2842
```

```
Qy   501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db  2843 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 2902

Qy   521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db  2903 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACC 2962

Qy   541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db  2963 ACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC 3022

Qy   561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db  3023 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATT 3082

Qy   581 TyrTyrSerLeuValMetSerGln------------------AspIleLysLysAlaAsp 594
Db  3083 TACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 3142

Qy   595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
Db  3143 GATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGT 3202

Qy   615 SerVal------------------LysSerIleLys---SerAspPheIleAspLysAla 627
Db  3203 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAG 3262

Qy   628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
Db  3263 CTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATG 3322

Qy   648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
Db  3323 AAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAAT 3382

Qy   668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
Db  3383 GGAGGCCTGCAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGA 3442

Qy   688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db  3443 ATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCC 3502

Qy   707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db  3503 CTTGGAATTATTTCTTTTATTACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGA 3562

Qy   727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db  3563 GAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTG 3622

Qy   747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
Db  3623 AGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGAT 3682

Qy   767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db  3683 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCA 3742

Qy   787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
Db  3743 AATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTC 3802

Qy   807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db  3803 TTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGA 3862

Qy   827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db  3863 CAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 3922

Qy   847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db  3923 GAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCT 3982

Qy   867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db  3983 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 4042

Qy   887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db  4043 TTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCC 4102

Qy   907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db  4103 TACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTT 4162

Qy   927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db  4163 GTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCC 4222

Qy   947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db  4223 AAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTAC 4282

Qy   967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db  4283 AGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTA 4342

Qy   987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db  4343 TTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAG 4402

Qy  1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
Db  4403 AAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAG 4462

Qy  1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
Db  4463 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATA 4522

Qy  1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
Db  4523 AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATC 4582

Qy  1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db  4583 CTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCA 4642

Qy  1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db  4643 CAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTG 4702

Qy  1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
Db  4703 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAA 4762

Qy  1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
Db  4763 CAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAA 4822

Qy  1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
Db  4823 GCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCC 4882

Qy  1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db  4883 AGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGAC 4942

Qy  1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
Db  4943 TTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTG 5002

Qy  1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
```

```
Db    5003 GCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 5044

RESULT 10
US-09-769-097-3
; Sequence 3, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
;  APPLICANT: Kimberly Anne Brun
;  APPLICANT:  Richard James Chenery
;  APPLICANT:  Harma Ellens
;  APPLICANT:  John Anthony Feild
;  APPLICANT:  Lin Yue
;  TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
;  TITLE OF INVENTION:  ENCODING RAT MDR1A AND SCREENING METHODS THEREOF
;  FILE REFERENCE: GP-50009-C2
;  CURRENT APPLICATION NUMBER: US/09/769,097
;  CURRENT FILING DATE:  2001-01-25
;  PRIOR APPLICATION NUMBER: 09/208,809
;  PRIOR FILING DATE: 1998-12-09
;  PRIOR APPLICATION NUMBER: 09/156,800
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: US99/20770
;  PRIOR FILING DATE: 1999-09-10
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 4425
;   TYPE: DNA
;   ORGANISM: RATTUS RATTUS
US-09-769-097-3

Alignment Scores:
Pred. No.:              9.44e-313     Length:          4425
Score:                  3517.00       Matches:         680
Percent Similarity:     73.90%        Conservative:    243
Best Local Similarity:  54.44%        Mismatches:      266
Query Match:            57.10%        Indels:          60
DB:                     10            Gaps:            10

US-09-873-409-5 (1-1222) x US-09-769-097-3 (1-4425)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db     499 ATGCTGCTGGGAACTCTGGCGGCCATTATCCATGGAATTGCGCTCCCACTTATGATGCTG 558

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db     559 GTCTTTGGAGACATGACAGATAGCTTTGCAAAT------GTAGGAAACAACCGTAGTATG 612

Qy      41 SerPhePheArgLeuThr------------------------------------------ 46
Db     613 AGTTTCTACAATGCTACAGACATATATGCCAAGCTGGAGGACGAAATGACCACGTACGCC 672

Qy      47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
Db     673 TACTATTACACGGGCATTGGTGCCGGTGTGCTCATCGTTGCCTACATCCAGGTTTCCACT 732

Qy      67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
Db     733 TGGTGCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAAGTTTTTCCATGCCATC 792

Qy      87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
Db     793 ATGAATCAGGAGATAGGCTGGTTTGACGTGCATGACGTTGGGGAGCTCAACACCCGGCTC 852

Qy     107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
Db     853 ACAGATGACGTCTCCAAAATTAATGAAGGAATTGGTGACAAAATTGGAATGTTCTTTCAG 912

Qy     126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
Db     913 GCAATGGCAACATTTTTTGGTGGTTTTATAATAGGATTTACTCGCGGCTGGAAGCTAACT 972

Qy     146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
Db     973 CTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTGTCAGCTGGTATTTGGGCAAAGATA 1032

Qy     166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
Db    1033 TTGTCTTCATTTACTGATAAGGAACTCCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAA 1092

Qy     186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
Db    1093 GAAGTCTTAGCAGCCATCAGAACTGTGATTGCCTTTGGAGGACAAAAGAAGGAACTTGAA 1152

Qy     206 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeu 225
Db    1152 ------------------------------------------------------------ 1152

Qy     226 SerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 245
Db    1153 ---------------------AGGTACAATAACAATTTGGAAGAAGCTAAAAGGCTTGGG 1191

Qy     246 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 265
Db    1192 ATAAAGAAAGCTATCACGGCCAACATTTCCATGGGTGCAGCTTTTCTGCTTATCTATGCA 1251

Qy     266 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 285
Db    1252 TCATATGCTCTGGCATTCTGGTATGGGACTTCCTTGGTCATCTCAAAAGAA------TAC 1305

Qy     286 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 305
Db    1306 ACTATTGGACAAGTGCTCACTGTCTTTTTTTCTGTATTAATTGGAGCATTCAGTGTTGGG 1365

Qy     306 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 325
Db    1366 CAGGCATCTCCAAATATTGAAGCCTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTC 1425

Qy     326 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 345
Db    1426 AGTATAATTGATAATAAGCCCAGTATAGACAGCTTCTCAAAGAGTGGGCACAAACCCGAC 1485

Qy     346 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 365
Db    1486 AACATACAAGGAAATTTGGAATTCAAAAATATTCACTTCAGTTACCCGTCTCGAAAAGAC 1545

Qy     366 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 385
Db    1546 GTTCAGATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTGGTT 1605

Qy     386 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 405
Db    1606 GGCAACAGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCC 1665

Qy     406 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 425
Db    1666 ATAGAGGGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTG 1725

Qy     426 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 445
Db    1726 CGGGAAATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCCACCACAATTGCCGAA 1785

Qy     446 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 465
Db    1786 AACATTCGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAA 1845

Qy     466 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 485
Db    1846 GCCAATGCCTATGATTTCATCATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAG 1905

Qy     486 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 505
Db    1906 AGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTC 1965

Qy     506 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 525
```

```
Db     1966 CGCAACCCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAA 2025

Qy      526 SerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 545
Db     2026 GCCGTGGTTCAGGCCGCTCTGGATAAGGCTAGAGAAGGCCGGACCACCATTGTGATAGCT 2085

Qy      546 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 565
Db     2086 CACCGCTTGTCTACAGTTCGCAATGCTGAYGTCATTGCTGGTTTTGATGGTGGTGTCATT 2145

Qy      566 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 585
Db     2146 GTGGAGCAAGGAAATCATGATGAGCTCATGAGAGAGAAAGGAATTTACTTCAAACTTGTC 2205

Qy      586 MetSerGln------------AspIleLysLysAlaAspGluGlnMetGluSer------ 599
Db     2206 ATGACTCAGACAGCAGGAAATGAAATTGAATTAGGAAATGAAGCTTGTGAATCTAAAGAY 2265

Qy      600 ------------MetThrTyrSerThrGluArgLysThrAsnSerLeu------ProLeu 613
Db     2266 GGAATTGATAATGTGGACATGTCTTCAAAAGATTCRGGATCCAGTCTAATAAGAAGAAGA 2325

Qy      614 HisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---Gln 632
Db     2326 TCAACTCGCAAAAGCATCCGTGGGCCACATGATCAAGACGGGGAACTTAGCACCAAAGAG 2385

Qy      633 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 652
Db     2386 GCTCTGGATGACGACGTACCTCCAGCTTCCTTTTGGCGGATCCTGAAGTTGAATTCAACT 2445

Qy      653 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 672
Db     2446 GAATGGCCTTATTTTGTGGTTGGTGTATTTTGTGCCATAATAAATGGAGGCTTGCAACCA 2505

Qy      673 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 692
Db     2506 GCATTCTCCATAATATTTTCAAAGGTTGTAGGGGTTTTTACAAAAAATGACACCCCTGAA 2565

Qy      693 Leu---LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 711
Db     2566 ATCCAGCGGCAGAACAGCAACTTGTTTTCTTTATTGTTCCTGATCCTTGGGATCATCTCT 2625

Qy      712 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 731
Db     2626 TTCATTACGTTTTTCCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAG 2685

Qy      732 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 751
Db     2686 CGACTCCGATACATGGTCTTCAAATCCATGCTGAGACAGGACATAAGCTGGTTTGATGAC 2745

Qy      752 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 771
Db     2746 CCTAAAAACACCACAGGAGCGCTGACCACCAGGCTTGCCAATGACGCTGCTCAAGTGAAA 2805

Qy      772 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 791
Db     2806 GGGGCTACAGGGTCTAGGCTTGCTGTTATTACCCAGAACATAGCAAATCTTGGGACAGGC 2865

Qy      792 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 811
Db     2866 ATCATCATATCCCTGATCTACGGCTGGCAATTGACACTTTTACTCCTAGCAATTGTTCCC 2925

Qy      812 ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp 831
Db     2926 ATCATTGCTATAGCAGGAGTGGTTGAAATGAAAATGTTGTCTGGACAAGCGCTGAAAGAT 2985

Qy      832 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 851
Db     2986 AAGAAGGAACTAGAAGGTTCTGGGAAGATCGCTACAGAAGCAATTGAAAACTTTCGCACT 3045

Qy      852 IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 871
Db     3046 GTCGTCTCTTTGACTCGGGAGCAGAAGTTTGAAACTATGTATGCCCAGAGCTTGCAGATA 3105

Qy      872 GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis 891
Db     3106 CCATACAGAAATGCTTTGAAGAAAGCGCACGTCTTTGGGATCACTTTCTCCTTCACCCAG 3165

Qy      892 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla 911
Db     3166 GCCATGATGTATTTCTCCTATGCTGCTTGTTTCCGGTTTGATGCCTACTTGGTGGCACGA 3225

Qy      912 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 931
Db     3226 GAACTCATGACATTTGAAAATGTTCTGTTAGTATTCTCAGCTATTGTCTTTGGTGCCATG 3285

Qy      932 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 951
Db     3286 GCAGTGGGGCAGGTCAGTTCATTCGCTCCTGACTACGCGAAAGCCAAAGTCTCGGCATCC 3345

Qy      952 HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys 971
Db     3346 CACATCATCAGGATCATTGAGAAAATCCCTGAGATTGACAGCTACAGCACGGAGGGCTTG 3405

Qy      972 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys 991
Db     3406 AAGCCTAATATGTTGGAAGGAAATGTGAAATTTAATGGAGTCATGTTCAACTATCCCACC 3465

Qy      992 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 1011
Db     3466 CGACCCAACATCCCAGTGCTTCAGGGGCTGAGCCTAGAGGTGAAGAAAGGGCAGACGCTG 3525

Qy     1012 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 1031
Db     3526 GCCCTCGTGGGCAGCAGTGGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTTGAGCGCTTC 3585

Qy     1032 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 1051
Db     3586 TATGACCCCATGGCCGGAACAGTGTTTCTAGATGGCAAAGAAATAAAGCAACTCAATGTC 3645

Qy     1052 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 1071
Db     3646 CAGTGGCTCCGCGCCCACCTGGGCATTGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGC 3705

Qy     1072 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 1091
Db     3706 ATCGCCGAGAACATTGCCTACGGAGACAACAGCCGTGTCGTGTCTCATAAGGAGATCGTG 3765

Qy     1092 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn 1111
Db     3766 AAGGCAGCCAAGGAGGCCAACATCCACCAGTTCATCGACTCACTGCCTGAGAAATACAAC 3825

Qy     1112 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 1131
Db     3826 ACCAGAGTGGGAGACAAAGGGACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATCGCCATC 3885

Qy     1132 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 1151
Db     3886 GCGCGCGCCCTCGTCAGACAGCCTCACATCTTACTTCTGGATGAAGCGACATCAGCTCTG 3945

Qy     1152 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1171
Db     3946 GATACGGAGAGTGAAAAGGTCGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACC 4005

Qy     1172 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1191
Db     4006 TGCATTGTGATCGCGCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATT 4065

Qy     1192 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1211
Db     4066 CAGAACGGCCAGGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCCCAGAAAGGCATC 4125

Qy     1212 TyrPheLysLeuValAsnAlaGlnSer 1220
Db     4126 TATTTCTCGATGGTCAGTGTGCAGGCT 4152
```

```
RESULT 11
US-09-769-097-1
; Sequence 1, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
;  APPLICANT: Kimberly Anne Brun
;  APPLICANT:  Richard James Chenery
;  APPLICANT:  Harma Ellens
;  APPLICANT:  John Anthony Feild
;  APPLICANT:  Lin Yue
;  TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
;  TITLE OF INVENTION:  ENCODING RAT MDR1A AND SCREENING METHODS THEREOF
;  FILE REFERENCE: GP-50009-C2
;  CURRENT APPLICATION NUMBER: US/09/769,097
;  CURRENT FILING DATE:  2001-01-25
;  PRIOR APPLICATION NUMBER: 09/208,809
;  PRIOR FILING DATE: 1998-12-09
;  PRIOR APPLICATION NUMBER: 09/156,800
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: US99/20770
;  PRIOR FILING DATE: 1999-09-10
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 4369
;   TYPE: DNA
;   ORGANISM: RATTUS RATTUS
US-09-769-097-1

Alignment Scores:
Pred. No.:               1.41e-312     Length:         4369
Score:                   3515.00       Matches:        680
Percent Similarity:      73.90%        Conservative:   243
Best Local Similarity:   54.44%        Mismatches:     266
Query Match:             57.07%        Indels:         60
DB:                      10            Gaps:           10

US-09-873-409-5 (1-1222) x US-09-769-097-1 (1-4369)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    499 ATGCTGCTGGGAACTCTGGCGGCCATTATCCATGGAATTGCGCTCCCACTTATGATGCTG 558

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db    559 GTCTTTGGAGACATGACAGATAGCTTTGCAAAT------GTAGGAAACAACCGTAGTATG 612

Qy     41 SerPhePheArgLeuThr------------------------------------------ 46
Db    613 AGTTTCTACAATGCTACAGACATATATGCCAAGCTGGAGGACGAAATGGCCACGTACGCC 672

Qy     47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
Db    673 TACTATTACACGGGCAATGGTGCCGGTGTGCTCATCGTTGCCTACATCCAGGTTTCACTT 732

Qy     67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
Db    733 TGGTGCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAAGTTTTTCCATGCCATC 792

Qy     87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
Db    793 ATGAATCAGGAGATAGGCTGGTTTGACGTGCATGACGTTGGGGAGCTCAACACCCGGCTC 852

Qy    107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
Db    853 ACAGATGACGTCTCCAAAATTAATGAAGGAATTGGTGACAAAATTGGAATGTTCTTTCAG 912

Qy    126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
Db    913 GCAATGGCAACATTTTTTGGTGGTTTTATAATAGGATTTACTCGCGGCTGGAAGCTAACT 972

Qy    146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
Db    973 CTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTGTCAGCTGGTATTTGGGCAAAGATA 1032

Qy    166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
Db   1033 TTGTCTTCATTTACTGATAAGGAACTCCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAA 1092

Qy    186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
Db   1093 GAAGTCTTAGCAGCCATCAGAACTGTGATTGCCTTTGGAGGACAAAAGAAGGAACTTGAA 1152

Qy    206 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeu 225
Db   1152 ------------------------------------------------------------ 1152

Qy    226 SerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 245
Db   1153 ---------------------AGGTACAATAACAATTTGGAAGAAGCTAAAAGGCTTGGG 1191

Qy    246 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 265
Db   1192 ATAAAGAAAGCTATCACGGCCAACATTTCCATGGGTGCAGCTTTTCTGCTTATCTATGCA 1251

Qy    266 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 285
Db   1252 TCATATGCTCTGGCATTCTGGTATGGGACTTCCTTGGTCATCTCAAAAGAA------TAC 1305

Qy    286 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 305
Db   1306 ACTATTGGACAAGTGCTCACTGTCTTTTTTTCTGTATTAATTGGAGCATTCAGTGTTGGG 1365

Qy    306 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 325
Db   1366 CAGGCATCTCCAAATATTGAAGCCTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTC 1425

Qy    326 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 345
Db   1426 AGTATAATTGATAATAAGCCCAGTATAGACAGCTTCTCAAAGAGTGGGCACAAACCCGAC 1485

Qy    346 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 365
Db   1486 AACATACAAGGAAATTTGGAATTCAAAAATATTCACTTCAGTTACCCGTCTCGAAAAGAC 1545

Qy    366 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 385
Db   1546 GTTCAGATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTGGTT 1605

Qy    386 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 405
Db   1606 GGCAACAGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCC 1665

Qy    406 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 425
Db   1666 ATAGAGGGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTG 1725

Qy    426 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 445
Db   1726 CGGGAAATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCCACCACAATTGCCGAA 1785

Qy    446 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 465
Db   1786 AACATTCGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAA 1845

Qy    466 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 485
Db   1846 GCCAATGCCTATGATTTCATCATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAG 1905

Qy    486 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 505
Db   1906 AGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTC 1965

Qy    506 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 525
```

```
Db   1966 CGCAACCCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAA 2025

Qy    526 SerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 545
Db   2026 GCCGTGGTTCAGGCCGCTCTGGATAAGGCTAGAGAAGGCCGGACCACCATTGTGATAGCT 2085

Qy    546 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 565
Db   2086 CACCGCTTGTCTACAGTTCGCAATGCTGACGTCATTGCTGGTTTGATGGTGGTGTCATT 2145

Qy    566 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 585
Db   2146 GTGGAGCAAGGAAATCATGATGAGCTCATGAGAGAGAAAGGAATTTACTTCAAACTTGTC 2205

Qy    586 MetSerGln------------AspIleLysLysAlaAspGluGlnMetGluSer------ 599
Db   2206 ATGACTCAGACAGCAGGAAATGAAATTGAATTAGGAAATGAAGCTTGTGAATCTAAAGAC 2265

Qy    600 ------------MetThrTyrSerThrGluArgLysThrAsnSerLeu------ProLeu 613
Db   2266 GGAATTGATAATGTGGACATGTCTTCAAAAGATTCGGGATCCAGTCTAATAAGAAGAAGA 2325

Qy    614 HisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---Gln 632
Db   2326 TCAACTCGCAAAAGCATCCGTGGGCCACATGATCAAGACGGGGAACTTAGCACCAAAGAG 2385

Qy    633 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 652
Db   2386 GCTCTGGATGACGACGTACCTCCAGCTTCCTTTTGGCGGATCCTGAAGTTGAATTCAACT 2445

Qy    653 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 672
Db   2446 GAATGGCCTTATTTTGTGGTTGGTGTATTTTGTGCCATAATAAATGGAGGCTTGCAACCA 2505

Qy    673 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 692
Db   2506 GCATTCTCCATAATATTTTCAAAGGTTGTAGGGGTTTTTACAAAAAATGACACCCCTGAA 2565

Qy    693 Leu---LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 711
Db   2566 ATCCAGCGGCAGAACAGCAACTTGTTTCTTTATTGTTCCTGATCCTTGGGATCATCTCT 2625

Qy    712 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 731
Db   2626 TTCATTACGTTTTTCCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAG 2685

Qy    732 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 751
Db   2686 CGACTCCGATACATGGTCTTCAAATCCATGCTGAGACAGGACATAAGCTGGTTTGATGAC 2745

Qy    752 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 771
Db   2746 CCTAAAAACACCACAGGAGCGCTGACCACCAGGCTTGCCAATGACGCTGCTCAAGTGAAA 2805

Qy    772 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 791
Db   2806 GGGGCTACAGGGTCTAGGCTTGCTGTTATTACCCAGAACATAGCAAATCTTGGGACAGGC 2865

Qy    792 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 811
Db   2866 ATCATCATATCCCTGATCTACGGCTGGCAATTGACACTTTTACTCCTAGCAATTGTTCCC 2925

Qy    812 ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp 831
Db   2926 ATCATTGCTATAGCAGGAGTGGTTGAAATGAAAATGTTGTCTGGACAAGCGCTGAAAGAT 2985

Qy    832 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 851
Db   2986 AAGAAGGAACTAGAAGGTTCTGGGAAGATCGCTACAGAAGCAATTGAAAACTTTCGCACT 3045

Qy    852 IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 871
Db   3046 GTCGTCTCTTTGACTCGGGAGCAGAAGTTTGAAACTATGTATGCCCAGAGCTTGCAGATA 3105

Qy    872 GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis 891
Db   3106 CCATACAGAAATGCTTTGAAGAAAGCGCACGTCTTTGGGATCACTTTCTCCTTCACCCAG 3165

Qy    892 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla 911
Db   3166 GCCATGATGTATTTCTCCTATGCTGCTTGTTTCCGGTTTGATGCCTACTTGGTGGCACGA 3225

Qy    912 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 931
Db   3226 GAACTCATGACATTTGAAAATGTTCTGTTAGTATTCTCAGCTATTGTCTTTGGTGCCATG 3285

Qy    932 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 951
Db   3286 GCAGTGGGGCAGGTCAGTTCATTCGCTCCTGACTACGCGAAAGCCAAAGTCTCGGCATCC 3345

Qy    952 HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys 971
Db   3346 CACATCATCAGGATCATTGAGAAAATCCCTGAGATTGACAGCTACAGCACGGAGGGCTTG 3405

Qy    972 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys 991
Db   3406 AAGCCTAATATGTTGGAAGGAAATGTGAAATTTAATGGAGTCATGTTCAACTATCCCACC 3465

Qy    992 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 1011
Db   3466 CGACCCAACATCCCAGTGCTTCAGGGGCTGAGCCTAGAGGTGAAGAAAGGGCAGACGCTG 3525

Qy   1012 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 1031
Db   3526 GCCCTCGTGGGCAGCAGTGGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTTGAGCGCTTC 3585

Qy   1032 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 1051
Db   3586 TATGACCCCATGGCCGGAACAGTGTTTCTAGATGGCAAAGAAATAAAGCAACTCAATGTC 3645

Qy   1052 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 1071
Db   3646 CAGTGGCTCCGCGCCCACCTGGGCATTGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGC 3705

Qy   1072 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 1091
Db   3706 ATCGCCGAGAACATTGCCTACGGAGACAACAGCCGTGTCGTGTCTCATAAGGAGATCGTG 3765

Qy   1092 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn 1111
Db   3766 AAGGCAGCCAAGGAGGCCAACATCCACCAGTTCATCGACTCACTGCCTGAGAAATACAAC 3825

Qy   1112 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 1131
Db   3826 ACCAGAGTGGGAGACAAAGGGACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATCGCCATC 3885

Qy   1132 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 1151
Db   3886 GCGCGCGCCCTCGTCAGACAGCCTCACATCTTACTTCTGGATGAAGCGACATCAGCTCTG 3945

Qy   1152 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1171
Db   3946 GATACGGAGAGTGAAAAGGTCGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACC 4005

Qy   1172 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1191
Db   4006 TGCATTGTGATCGCGCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATT 4065

Qy   1192 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1211
Db   4066 CAGAACGGCCAGGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCCCAGAAAGGCATC 4125

Qy   1212 TyrPheLysLeuValAsnAlaGlnSer 1220
Db   4126 TATTTCTCGATGGTCAGTGTGCAGGCT 4152
```

```
RESULT 12
US-10-072-621-2
; Sequence 2, Application US/10072621
; Patent No. US20020169137A1
; GENERAL INFORMATION:
;  APPLICANT: Reiner, Peter B.
;  APPLICANT:  Connop, Bruce P.
;  APPLICANT:  Pollard, Michelle
;  TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION
;  TITLE OF INVENTION:  BY MODIFICATION OF ABC TRANSPORTER EXPRESSION OR ACTIVITY
;  FILE REFERENCE: 100103.402
;  CURRENT APPLICATION NUMBER: US/10/072,621
;  CURRENT FILING DATE:  2002-02-08
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 4643
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-072-621-2

Alignment Scores:
Pred. No.:              1.91e-312    Length:         4643
Score:                  3514.00      Matches:        681
Percent Similarity:     73.92%       Conservative:   246
Best Local Similarity:  54.31%       Mismatches:     263
Query Match:            57.05%       Indels:         64
DB:                     9            Gaps:           11

US-09-873-409-5 (1-1222) x US-10-072-621-2 (1-4643)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    575 ATGGTGGTGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy     21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db    635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe----------------- 43
Db    695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    875 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    935 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAAATT 994

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db   1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db   1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db   1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db   1235 AAGAAAGAACTTGAA--------------------------------------------- 1249

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db   1250 -----------------------------------AGGTACAACAAAAATTTAGAAGAA 1273

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db   1274 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 1333

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db   1334 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 1393

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db   1394 GGGGAA------TATTCTATTGGACAAGTACTCACTGTA---TTTTCTGTATTAATTGGG 1444

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db   1445 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1504

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   1505 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1564

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   1565 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1624

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   1625 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGTCTGAACCTGAAGGTGCAGAGTGGGCAG 1684

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   1685 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1744

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   1745 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1804

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   1805 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1864

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   1865 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1924

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   1925 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1984

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   1985 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 2044

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   2045 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 2104

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db   2105 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACC 2164

Qy    541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db   2165 ACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC 2224
```

```
Qy    561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db   2225 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATT 2284

Qy    581 TyrTyrSerLeuValMetSerGln------------------AspIleLysLysAlaAsp 594
Db   2285 TACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 2344

Qy    595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
Db   2345 GATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGT 2404

Qy    615 SerVal------------------LysSerIleLys---SerAspPheIleAspLysAla 627
Db   2405 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAG 2464

Qy    628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
Db   2465 CTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATG 2524

Qy    648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
Db   2525 AAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAAT 2584

Qy    668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
Db   2585 GGAGGCCTGCAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGA 2644

Qy    688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db   2645 ATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCC 2704

Qy    707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db   2705 CTTGGAATTATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGA 2764

Qy    727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db   2765 GAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTG 2824

Qy    747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
Db   2825 AGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGAT 2884

Qy    767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db   2885 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCA 2944

Qy    787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
Db   2945 AATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTC 3004

Qy    807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db   3005 TTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGA 3064

Qy    827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db   3065 CAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 3124

Qy    847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db   3125 GAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCT 3184

Qy    867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db   3185 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 3244

Qy    887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db   3245 TTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCC 3304

Qy    907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db   3305 TACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTT 3364

Qy    927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db   3365 GTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCC 3424

Qy    947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db   3425 AAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTAC 3484

Qy    967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db   3485 AGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTA 3544

Qy    987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db   3545 TTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAG 3604

Qy   1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
Db   3605 AAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAG 3664

Qy   1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
Db   3665 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATA 3724

Qy   1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
Db   3725 AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATC 3784

Qy   1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db   3785 CTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCA 3844

Qy   1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db   3845 CAGGAAGAGATTGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTG 3904

Qy   1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
Db   3905 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAA 3964

Qy   1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
Db   3965 CAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAA 4024

Qy   1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
Db   4025 GCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCC 4084

Qy   1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db   4085 AGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGAC 4144

Qy   1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
Db   4145 TTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTG 4204

Qy   1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db   4205 GCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 4246
```

RESULT 13
US-09-866-866A-5
; Sequence 5, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2

```
;  CURRENT APPLICATION NUMBER: US/09/866,866A
;  CURRENT FILING DATE:  2001-08-30
;  PRIOR APPLICATION NUMBER: 09/584,586
;  PRIOR FILING DATE: `2000-05-31
;  PRIOR APPLICATION NUMBER: PCT/US99/11825
;  PRIOR FILING DATE: 1999-05-27
;  PRIOR APPLICATION NUMBER: 60/086,988
;  PRIOR FILING DATE: 1998-05-28
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 4189
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-866-866A-5

Alignment Scores:
Pred. No.:              1.68e-311     Length:        4189
Score:                  3503.00       Matches:       678
Percent Similarity:     73.13%        Conservative:  242
Best Local Similarity:  53.90%        Mismatches:    262
Query Match:            56.88%        Indels:        76
DB:                     10            Gaps:          11

US-09-873-409-5 (1-1222) x US-09-866-866A-5 (1-4189)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           ||||||||||||   ||||||::::::::::::|||      ||||||||||:::     |||
Db     148 ATGATTCTGGGAACTCTCGCTGCTATTATCCATGGAACATTACTTCCCCTCTTGATGCTG 207

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
           |||   |||   |||:::|||:::
Db     208 GTGTTTGGAAACATGACAGATAGTTTACAAAAGCAGAAGCCAGTATTCTGCCAAGCATT 267

Qy      31 ------------------SerGlyCysLeuValGlnThrAsnThr--------------- 39
                             :::      |||:::   :::|||:::
Db     268 ACTAATCAAAGTGGACCCAACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAAGAG 327

Qy      40 ---------TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56
                    |||::::::                   ||||||   |||||||||
Db     328 ATGGCCATATACGCCTAC---------------TATTACACCGGGATTGGTGCTGGTGTG 372

Qy      57 LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys 76
           ||||||       |||||||||:::|||||||||    :::   |||    ||||||
Db     373 CTCATAGTTGCCTACATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGAAGACAGATACAC 432

Qy      77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96
           :::|||||||||::::::||||||||:::::::::||||||||||||||||
Db     433 AAGATTAGGCAGAAGTTTTTCCATGCTATAATGAATCAGGAGATAGGCTGGTTTGATGTG 492

Qy      97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
              |||:::||||||||||||||||||:::|||   |||:::   ||||||:::||||||
Db     493 CATGATGTTGGGGAGCTCAACACCCGGCTCACAGATGATGTCTCCAAAATTAATGACGGA 552

Qy     116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
           |||||||||||||||   :::   ||||||:::::::::||||||       |||
Db     553 ATTGGTGACAAAATTGGGATGTTTTTCAGTCCATAACCACATTTTAGCCGGTTTTATC 612

Qy     136 ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
           :::|||   :::   ||||||||||||||||||||||||   |||:::   ||||||||||||
Db     613 ATAGGATTTATAAGTGGTTGGAAGCTAACCCTTGTCATTTTGGCTGTCAGCCCTCTTATT 672

Qy     156 MetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
              |||:::|||      ::::::::::::   |||   |||:::|||||||||
Db     673 GGATTGTCATCTGCTTTGTGGGCAAAGGTATTGACTTCATTTACTAATAAGGAACTCCAG 732

Qy     176 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIle 195
           ||||||:::||||||||||||||||||||||||||||||::::::||||||||||||||||
Db     733 GCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT 792

Qy     196 AlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyr 215
           ||||||      |||:::|||||||||:::
Db     793 GCCTTTGGAGGACAACAGAAGGAACTTGAA------------------------------ 822

Qy     216 AlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThr 235
                                                            ||||||
Db     823 ---------------------------------------------------AGGTACAAT 831

Qy     236 GlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSer 255
           :::||||||::::::||||||:::   |||||||||:::   |||   :::   :::|||
Db     832 AAAAATTTAGAAGAAGCTAAAAATGTTGGCATAAAGAAAGCTATCACAGCCAGCATTTCG 891

Qy     256 LeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThr 275
           :::|||      |||      :::      :::|||   ||||||||||||||||||||||
Db     892 ATAGGCATTGCCTACCTGTTGGTCTATGCATCATATGCACTGGCATTCTGGTATGGGACA 951

Qy     276 SerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePhe 295
           ||||||:::|||:::   |||      |||:::||||||   ||||||   |||||||||
Db     952 TCCTTGGTCCTCTCAAATGAA------TATTCTATTGGAGAAGTGCTTACTGTCTTCTTC 1005

Qy     296 SerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAla 315
           |||::::::      ::::::   ||||||         |||:::   |||   ||||||
Db    1006 TCTATTTTGTTGGGGACTTTTAGTATTGGACACTTGGCCCCAAACATAGAAGCCTTTGCA 1065

Qy     316 IleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAsp 335
              ||||||||||||||||||   ||||||::::::||||||   :::||||||||||||
Db    1066 AACGCACGAGGGGCAGCCTTTGAAATCTTCAAGATAATTGATAACGAGCCAAGCATTGAC 1125

Qy     336 AsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsn 355
           :::|||||||||   ||||||||||||:::||||||   |||   :::||||||||||||
Db    1126 AGCTTCTCAACAAAGGGCTACAAACCAGACAGTATAATGGGAAACTTAGAGTTTAAAAAT 1185

Qy     356 ValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArg 375
           |||   ||||||||||||||||||       ::::::||||||||||||||||||:::
Db    1186 GTTCACTTCAACTACCCATCGAGAAGCGAAGTTCAGATCTTGAAGGGCCTCAATCTGAAG 1245

Qy     376 IleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVal 395
           :::||||||:::||||||||||||||||||      :::|||   ||||||||||||
Db    1246 GTGAAGAGCGGACAGACGGTGGCCTTGGTTGGCAACAGTGGCTGTGGAAAAAGCACAACT 1305

Qy     396 ValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsn 415
           ||||||||:::||||||||||||||||||      :::|||   :::   :::|||
Db    1306 GTCCAGCTGATGCAGAGGCTCTACGACCCCCTGGAGGGCGTGGTCAGTATCGACGGACAA 1365

Qy     416 AspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGlu 435
           ||||||||   :::||||||||:::   |||:::   ||||||||||||||||||||||
Db    1366 GACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCATTGGTGTGGTGAGTCAGGAA 1425

Qy     436 ProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThr 455
           ||||||||||||   ||||||||:::   |||||:::||||||||||:::|||||||
Db    1426 CCTGTGCTGTTTGCCACCACGATCGCCGAGAACATTCGCTATGGCCGAGAAGATGTCACC 1485

Qy     456 AspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhe 475
           :::|||:::|||:::|||   :::|||||||||||||||||||||||||||:::
Db    1486 ATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTCATCATGAAACTG 1545

Qy     476 ProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLys 495
           |||::::::|||:::||||||||||||||||:::||||||||||:::||||||||||||
Db    1546 CCCCACCAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAA 1605

Qy     496 GlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGlu 515
           ||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db    1606 CAGAGAATCGCCATTGCCCGGGCCCTGGTCCGCAATCCCAAGATCCTTTTGTTGGACGAG 1665

Qy     516 AlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAla 535
           ||||||||||||||||||:::||||||::::::   ||||||||||||||:::||||||
Db    1666 GCCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTGGTGCAGGCCGCACTGGATAAGGCT 1725

Qy     536 SerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAsp 555
              :::|||||||||||||||||:::|||||||||||||||||||:::|||:::||||||
Db    1726 AGAGAAGGCCGGACCACCATTGTGATAGCTCATCGCTTGTCTACAGTTCGTAATGCTGAC 1785

Qy     556 LeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMet 575
           :::|||                  |||::::::   |||:::|||   |||   |||||||||
Db    1786 GTCATTGCTGGTTTTGATGGTGGTGTCATTGTGGAGCAAGGAAATCATGATGAGCTCATG 1845
```

```
Qy    576 AlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGlu 595
Db   1846 AGAGAAAAGGGCATTTACTTCAAACTTGTCATGACACAG---ACTAGAGGAAATGAAATT 1902

Qy    596 GlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSer 615
Db   1903 GAACCAGGAAATAATGCTTATGGATCCCAGAGTGACACTGATGCTTCTGAACTGACTTCA 1962

Qy    616 ValLysSerIleLysSerAspPheIleAspLysAla-----------------GluGlu 629
Db   1963 GAAGAATCC---AAATCACCTTTAATAAGGAGATCAATTTACAGAAGTGTCCACAGAAAG 2019

Qy    630 SerThrGlnSerLysGluIleSerLeu-----------------------ProGluVal 641
Db   2020 CAAGACCAAGAGAGAAGACTTAGTATGAAAGAGGCTGTGGATGAAGATGTGCCTCTGGTT 2079

Qy    642 SerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThr 661
Db   2080 TCCTTTTGGCGGATCCTAAATCTAAATCTAAGTGAATGGCCTTATTTACTTGTTGGCGTA 2139

Qy    662 LeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIle 681
Db   2140 CTTTGCGCTGTTATAAATGGGTGCATACAACCAGTGTTTGCCATAGTATTTTCAAGGATT 2199

Qy    682 IleThrMetPheGly---AsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 700
Db   2200 GTAGGGGTTTTTTCAAGAGATGATGACCATGAAACTAAACGACAGAATTGTAATTTGTTT 2259

Qy    701 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 720
Db   2260 TCCCTGTTCTTTCTGGTTATGGGGCTGATTTCTTTTGTTACATATTTCTTTCAGGGCTTC 2319

Qy    721 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 740
Db   2320 ACATTTGGCAAAGCCGGAGAGATCCTCACCAAGCGAGTCCGATACATGGTTTTCAAATCC 2379

Qy    741 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 760
Db   2380 ATGCTGAGACAGGATATAAGCTGGTTCGATGACCATAAGAACAGCACTGGCTCACTGACC 2439

Qy    761 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 780
Db   2440 ACCAGGCTCGCCAGTGATGCTTCTAGTGTTAAAGGGGCGATGGGCGCCAGGCTTGCTGTA 2499

Qy    781 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrp 800
Db   2500 GTTACCCAGAATGTAGCAAACCTCGGGACAGGAGTCATCCTCTCCTTAGTCTATGGCTGG 2559

Qy    801 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 820
Db   2560 CAGCTGACACTTCTACTTGTAGTAATTATACCGCTCATTGTATTGGGCGGAATTATTGAA 2619

Qy    821 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 840
Db   2620 ATGAAGCTGTTGTCTGGCCAAGCCTTGAAGGACAAGAAACAGCTTGAGATCTCTGGGAAG 2679

Qy    841 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 860
Db   2680 ATTGCTACAGAAGCAATTGAAAACTTCCGCACTATTGTCTCTTTGACTCGGGAGCAGAAG 2739

Qy    861 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 880
Db   2740 TTTGAAACCATGTATGCCCAGAGCTTGCAGGTACCATACAGAAATGCGATGAAGAAAGCA 2799

Qy    881 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 900
Db   2800 CACGTGTTTGGGATCACGTTCTCCTTCACCCAGGCCATGATGTATTTTTCTTATGCTGCT 2859

Qy    901 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 920
Db   2860 TGTTTCCGGTTCGGTGCCTACTTGGTGGCACAACAACTCATGACTTTTGAAAATGTTATG 2919

Qy    921 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 940
Db   2920 TTGGTATTTTCTGCTGTTGTCTTTGGTGCCATGGCAGCTGGGAATACTAGTTCATTTGCT 2979

Qy    941 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 960
Db   2980 CCTGACTATGCGAAAGCCAAAGTATCAGCATCTCATATCATCAGGATCATTGAGAAAACC 3039

Qy    961 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 980
Db   3040 CCTGAGATTGACAGCTACAGCACAGAGGGCTTGAAGCCTACTCTGTTAGAAGGAAATGTA 3099

Qy    981 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 1000
Db   3100 AAATTTAATGGAGTCCAGTTTAACTATCCCACCCGACCCAACATCCCAGTGCTTCAGGGG 3159

Qy   1001 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 1020
Db   3160 CTGAGCCTCGAGGTGAAGAAGGGCCAGACGTTGGCCCTGGTGGGCAGCAGTGGCTGTGGG 3219

Qy   1021 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 1040
Db   3220 AAGAGCACAGTGGTCCAGCTGCTCGAGCGCTTCTACGACCCCATGGCTGGATCAGTGTTT 3279

Qy   1041 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 1060
Db   3280 CTAGATGGCAAAGAAATAAAGCAACTGAATGTCCAGTGGCTCCGAGCTCACCTTGGCATT 3339

Qy   1061 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 1080
Db   3340 GTGTCCCAGGAGCCCATTCTCTTTGACTGCAGCATTGCAGAGAACATCGCCTATGGAGAC 3399

Qy   1081 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 1100
Db   3400 AACAGCCGGGCCGTGTCTCATGAGGAGATTGTGAGGGCAGCCAAGGAGGCCAACATCCAC 3459

Qy   1101 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 1120
Db   3460 CAGTTCATCGACTCACTGCCTGATAAATACAACACCAGAGTAGGAGACAAAGGCACTCAG 3519

Qy   1121 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 1140
Db   3520 CTGTCGGGTGGGCAGAAGCAGCGCATCGCCATCGCACGTGCCCTCGTCAGACAGCCTCAC 3579

Qy   1141 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 1160
Db   3580 ATTTACTTCTGGACGAAGCAACATCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAG 3639

Qy   1161 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 1180
Db   3640 GAAGCGCTGGACAAAGCCAGGGAAGGCCGCACCTGCATTGTGATCGCTCACCGCCTGTCC 3699

Qy   1181 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 1200
Db   3700 ACCATCCAGAACGCGGACTTGATCGTGGTGATTGAGAACGGCAAAGTCAAGGAGCACGGC 3759

Qy   1201 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 1218
Db   3760 ACCCACCAGCAGCTGCTGGCGCAGAAGGGCATCTACTTCTCAATGGTCCAGGCT 3813
```

RESULT 14
US-09-866-866A-7
; Sequence 7, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825

```
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
;   LENGTH: 4788
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-866-866A-7

Alignment Scores:
Pred. No.:              4.77e-311    Length:         4788
Score: .                3499.00      Matches:        668
Percent Similarity:     73.13%       Conservative:   252
Best Local Similarity:  53.10%       Mismatches:     264
Query Match:            56.81%       Indels:         74
DB:                     10           Gaps:           9

US-09-873-409-5 (1-1222) x US-09-866-866A-7 (1-4788)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    148 ATGCTGGTGGGAACTCTGGCTGCTATTATCCATGGAGTGGCGCTCCCACTTATGATGCTG 207

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsn------ 38
Db    208 ATCTTTGGTGACATGACAGATAGCTTTGCAAGTGTAGGAAACGTCTCTAAAAACAGTACT 267

Qy     39 ---------------------------------------------------------Thr 39
Db    268 AATATGAGTGAGGCCGATAAAAGAGCCATGTTTGCCAAACTGGAGGAAGAAATGACCACG 327

Qy     40 TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db    328 TACGCCTAC---------------TATTACACCGGGATTGGTGCTGGTGTGCTCATAGTT 372

Qy     60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db    373 GCCTACATCCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAGATACACAAGATCAGG 432

Qy     80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db    433 CAGAAGTTTTTCATGCTATAATGAATCAGGAGATAGGCTGGTTTGATGTGCATGACGTT 492

Qy    100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db    493 GGGGAGCTCAACACCCGGCTCACAGATGATGTTTCCAAAATTAATGAAGGAATTGGTGAC 552

Qy    119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db    553 AAAATCGGAATGTTCTTCCAGGCAATGGCAACATTTTTTGGTGGTTTTATAATAGGATTT 612

Qy    139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db    613 ACCCGTGGCTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTGTCA 672

Qy    159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db    673 GCTGGTATTTGGGCAAAGATATTGTCTTCATTTACTGATAAGGAACTCCATGCTTATGCA 732

Qy    179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db    733 AAAGCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCGTTTGGA 792

Qy    199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
Db    793 GGACAAAAGAAGGAACTTGAA--------------------------------------- 813

Qy    219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
Db    814 ------------------------------------------AGGTACAATAACAACTTG 831

Qy    239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
Db    832 GAAGAAGCTAAAAGGCTGGGGATAAAGAAAGCTATCACGGCCAACATCTCCATGGGTGCA 891

Qy    259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
Db    892 GCTTTTCTCCTTATCTATGCATCATATGCTCTGGCATTCTGGTATGGGACTTCCTTGGTC 951

Qy    279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
Db    952 ATCTCCAAAGAA------TACTCTATTGGACAAGTGCTCACTGTCTTCTTTTCCGTGTTA 1005

Qy    299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
Db   1006 ATTGGAGCATTCAGTGTTGGACAGGCATCTCCAAATATTGAAGCCTTCGCCAATGCACGA 1065

Qy    319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
Db   1066 GGAGCAGCTTATGAAGTCTTCAAAATAATTGATAATAAGCCCAGTATAGACAGCTTCTCA 1125

Qy    339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
Db   1126 AAGAGTGGGCACAAACCAGACAACATACAAGGAAATCTGGAATTTAAGAATATTCACTTC 1185

Qy    359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
Db   1186 AGTTACCCATCTCGAAAAGAAGTTCAGATCTTGAAGGGCCTCAATCTGAAGGTGAAGAGC 1245

Qy    379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
Db   1246 GGACAGACGGTGGCCCTGGTTGGCAACAGTGGCTGTGGAAAAAGCACAACTGTCCAGCTG 1305

Qy    399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
Db   1306 ATGCAAAGGCTCTACGACCCCCTAGATGGCATGGTCAGTATCGACGGACAGGACATCAGA 1365

Qy    419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db   1366 ACCATCAATGTGAGGTATCTGAGGGAGATCATTGGTGTGGTGAGTCAGGAACCTGTGCTG 1425

Qy    439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
Db   1426 TTTGCCACCACGATCGCCGAGAACATTCGCTATGGCCGAGAAGATGTCACCATGGATGAG 1485

Qy    459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
Db   1486 ATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTCATCATGAAACTGCCCCACCAA 1545

Qy    479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db   1546 TTTGACACCCTGGTTGGTGAGAGAGGGGCGCACGTGAGTGGGGGACAGAAACAGAGAATC 1605

Qy    499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db   1606 GCCATTGCCCGGGCCCTGGTCCGCAATCCCAAGATCCTTTTGTTGGACGAGGCCACCTCA 1665

Qy    519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 538
Db   1666 GCCCTGGATACAGAAAGTGAAGCTGTGGTTCAGGCCGCACTGGATAAGGCTAGAGAAGGC 1725

Qy    539 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558
Db   1726 CGGACCACCATTGTGATAGCTCATCGCTTGTCTACCGTTCGTAATGCTGACGTCATTGCT 1785

Qy    559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578
Db   1786 GGTTTTGATGGTGGTGTCATTGTGGAGCAAGGAAATCATGATGAGCTCATGAGAGAAAAG 1845

Qy    579 GlyLeuTyrTyrSerLeuValMetSerGln------------------------------ 588
Db   1846 GGCATTTACTTCAAACTTGTCATGACACAGACAGCAGGAAATGAAATTGAATTAGGAAAT 1905

Qy    589 AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThr 608
```

```
Db    1906 GAAGCTTGTAAATCTAAGGATGAAATTGATAATTTAGACATGTCTTCAAAAGATTCAGGA 1965

Qy     609 AsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle------------ 624
Db    1966 TCCAGTCTAATAAGAAGA------AGATCAACTCGCAAAAGCATCTGTGGACCACATGAC 2019

Qy     625 ---AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643
Db    2020 CAAGACAGGAAGCTTAGTACCAAAGAGGCCCTGGATGAAGATGTACCTCCAGCTTCCTTT 2079

Qy     644 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 663
Db    2080 TGGCGGATCCTGAAGTTGAATTCAACTGAATGGCCTTATTTTGTGGTTGGTATATTCTGT 2139

Qy     664 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 683
Db    2140 GCCATAATAAATGGAGGCTTACAGCCAGCATTCTCCGTAATATTTTCAAAAGTTGTAGGG 2199

Qy     684 MetPheGlyAsn---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 702
Db    2200 GTTTTTACAAATGGTGGCCCCCCTGAAACCCAGCGGCAGAACAGCAACTTGTTTTCCTTG 2259

Qy     703 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722
Db    2260 TTGTTTCTGATCCTTGGGATCATTTCTTTCATTACATTTTTTCTTCAGGGCTTCACATTT 2319

Qy     723 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 742
Db    2320 GGCAAAGCTGGAGAGATCCTCACCAAGCGACTCCGATACATGGTTTTCAAATCCATGCTG 2379

Qy     743 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIle 762
Db    2380 AGACAGGATGTGAGCTGGTTTGATGACCCTAAAAACACCACCGGAGCACTGACCACCAGG 2439

Qy     763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 782
Db    2440 CTCGCCAACGATGCTGCTCAAGTGAAAGGGGCTACAGGGTCTAGGCTTGCTGTGATTTTC 2499

Qy     783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 802
Db    2500 CAGAACATAGCAAATCTTGGGACAGGAATCATCATATCCCTAATCTATGGCTGGCAACTA 2559

Qy     803 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 822
Db    2560 ACACTTTTACTCTTAGCAATTGTACCCATCATTGCGATAGCTGGAGTGGTTGAAATGAAA 2619

Qy     823 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 842
Db    2620 ATGTTGTCTGGACAAGCACTGAAAGATAAGAAGGAACTAGAAGGTTCTGGAAAGATTGCT 2679

Qy     843 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 862
Db    2680 ACGGAAGCAATTGAAAACTTCCGCACTGTTGTCTCTTTGACTCGGGAGCAGAAGTTTGAA 2739

Qy     863 GlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 882
Db    2740 ACCATGTATGCCCAGAGCTTGCAGATACCATACAGAAATGCGATGAAGAAAGCACACGTG 2799

Qy     883 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 902
Db    2800 TTTGGGATCACGTTCTCCTTCACCCAGGCCATGATGTATTTTTCTTATGCTGCTTGTTTC 2859

Qy     903 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 922
Db    2860 CGGTTCGGTGCCTACTTGGTGACACAACAACTCATGACTTTTGAAAATGTTCTGTTAGTA 2919

Qy     923 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 942
Db    2920 TTCTCAGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAGGTCAGTTCATTCGCTCCTGAC 2979

Qy     943 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 962
Db    2980 TATGCGAAAGCAACAGTGTCAGCATCCCACATCATCAGGATCATTGAGAAAACCCCCGAG 3039

Qy     963 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 982
Db    3040 ATTGACAGCTACAGCACGCAAGGCCTAAAGCCGAATATGTTGGAAGGAAATGTGCAATTT 3099

Qy     983 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002
Db    3100 AGTGGAGTCGTGTTCAACTATCCCACCCGACCCAGCATCCCAGTGCTTCAGGGGCTGAGC 3159

Qy    1003 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 1022
Db    3160 CTTGAGGTGAAGAAGGGCCAGACGCTGGCCCTGGTGGGCAGCAGTGGCTGCGGGAAGAGC 3219

Qy    1023 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1042
Db    3220 ACAGTGGTCCAGCTGCTCGAGCGCTTCTACGACCCCATGGCTGGATCAGTGTTTCTAGAT 3279

Qy    1043 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 1062
Db    3280 GGCAAAGAAATAAAGCAACTGAATGTCCAGTGGCTCCGAGCACAGCTGGGCATTGTGTCC 3339

Qy    1063 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082
Db    3340 CAAGAGCCCATTCTCTTTGACTGCAGCATCGCAGAGAACATTGCCTACGGAGACAACAGC 3399

Qy    1083 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1102
Db    3400 CGGGTCGTGTCTTATGAGGAGATTGTGAGGGCAGCCAAGGAGGCCAACATCCACCAGTTC 3459

Qy    1103 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122
Db    3460 ATCGACTCGCTACCTGATAAATACAACACCAGAGTAGGAGACAAAGGCACTCAGCTGTCG 3519

Qy    1123 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 1142
Db    3520 GGTGGGCAGAAGCAGCGCATCGCCATCGCACGCGCCCTCGTCAGACAGCCTCACATTTTA 3579

Qy    1143 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 1162
Db    3580 CTTCTGGACGAAGCAACATCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAGGAAGCG 3639

Qy    1163 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 1182
Db    3640 CTGGACAAAGCCAGGGAAGGCCGCACCTGCATTGTGATCGCTCACCGCCTGTCCACCATC 3699

Qy    1183 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 1202
Db    3700 CAGAACGCGGACTTGATCGTGGTGATTCAGAACGGCAAGGTCAAGGAGCACGGCACCCAC 3759

Qy    1203 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db    3760 CAGCAGCTGCTGGCGCAGAAGGGCATCTACTTCTCAATGGTCAGTGTGCAGGCT 3813

RESULT 15
US-09-880-107-2299
; Sequence 2299, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, Darci T.
;  APPLICANT:  Vockley, Joseph G.
;  APPLICANT:  Scherf, Uwe
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
;  FILE REFERENCE: 44921-5028-WO
;  CURRENT APPLICATION NUMBER: US/09/880,107
;  CURRENT FILING DATE:  2001-06-14
;  PRIOR APPLICATION NUMBER: US 60/211,379
;  PRIOR FILING DATE: 2000-06-14
;  PRIOR APPLICATION NUMBER: US 60/237,054
;  PRIOR FILING DATE: 2000-10-02
;  NUMBER OF SEQ ID NOS: 3950
;  SOFTWARE: PatentIn Ver. 2.1
```

```
; SEQ ID NO 2299
;   LENGTH: 3924
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 M23234
US-09-880-107-2299

Alignment Scores:
Pred. No.:               1.59e-309    Length:        3924
Score:                   3481.00      Matches:       676
Percent Similarity:      72.82%       Conservative:  235
Best Local Similarity:   54.04%       Mismatches:    276
Query Match:             56.52%       Indels:        64
DB:                      10           Gaps:          9

US-09-873-409-5 (1-1222) x US-09-880-107-2299 (1-3924)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            |||   ||||||   :::   ::::::   :::|||:::   ||||||||||    :::
Db    201 ATGTCGCTGGGTACCATCATGGCCATAGCTCACGGATCAGGTCTCCCCCTCATGATGATA 260

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
            |||   ||||||||:::|||         :::                         :::
Db    261 GTATTTGGAGAGATGACTGACAAATTTGTTGAT-----------ACTGCAGGAAACTTC 308

Qy     41 SerPhe---------PheArgLeuThrLeu----------------------------- 47
          |||||||         |||   |||:::|||
Db    309 TCCTTTCCAGTGAACTTTTCCTTGTCGCTGCTAAATCCAGGCAAAATTCTGGAAGAAGAA 368

Qy     48 -----------------TyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                           ||||||   |||:::|||         |||:::       |||
Db    369 ATGACTAGATATGCATATTACTACTCAGGATTGGGTGCTGGAGTTCTTGTTGCTGCCTAT 428

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
          ||||||:::|||   |||   :::   |||   ||||||   ::::::||||||::::::
Db    429 ATACAAGTTTCATTTTGGACTTTGGCAGCTGGTCGACAGATCAGGAAAATTAGGCAGAAG 488

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
          ||||||||::::::|||   |||:::||||||||||||||||      |||      |||
Db    489 TTTTTCATGCTATTCTACGACAGGAAATAGGATGGTTTGACATCAATGACACCACTGAA 548

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
          ||||||||||||:::|||   ||||||   ||||||||:::||||||||||||||||:::
Db    549 CTCAATACGCGGCTAACAGATGACATCTCCAAAATCAGTGAAGGAATTGGTGACAAGGTT 608

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
            :::   ||||||   ::::::|||         ||||||         ::::::
Db    609 GGAATGTTCTTTCAAGCAGTAGCCACGTTTTTGCAGGATTCATAGTGGGATTCATCAGA 668

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
          ||||||||||||||||||||||   ::::::   ||||||::::::      |||||||||
Db    669 GGATGGAAGCTCACCCTTGTGATAATGGCCATCAGCCCTATTCTAGGACTCTCTGCAGCC 728

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
                ::::::::::::   :::   :::   |||||||||:::||||||:::||||||
Db    729 GTTTGGGCAAAGATACTCTCGGCATTTAGTGACAAAGAACTAGCTGCTTATGCAAAAGCA 788

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
          ||||||||||||||||||   |||   :::||||||||||||||||||||||    |||
Db    789 GGCGCCGTGGCAGAAGAGGCTCTGGGGGCCATCAGGACTGTGATAGCTTTCGGGGGCCAG 848

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
             |||||||||:::
Db    849 AACAAAGAGCTGGAA--------------------------------------------- 863

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                              ||||||   ::::::|||::::::
Db    864 -----------------------------------AGGTATCAGAAACATTTAGAAAAT 887

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
          ||||||:::   |||||||||:::   |||::::::   :::|||:::|||       :::
Db    888 GCCAAAGAGATTGGAATTAAAAAAGCTATTTCAGCAAACATTTCCATGGGTATTGCCTTC 947

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
                :::      :::|||   ||||||||||||||||||::::::|||:::::::::
Db    948 CTGTTAATATATGCATCATATGCACTGGCCTTCTGGTATGGATCCACTCTAGTCATATCA 1007

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
             |||      ||||||||||||         :::   ||||||||||||::::::
Db   1008 AAAGAA------TATACTATTGGAAATGCAATGACAGTTTTTTTTCAATCCTAATTGGA 1061

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
          ::::::   :::|||   |||   |||         :::   ||||||   |||||||||||
Db   1062 GCTTTCAGTGTTGGCCAGGCTGCCCCATGTATTGATGCTTTTGCCAATGCAAGAGGAGCA 1121

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
          |||:::   ||||||   :::||||||         |||   ||||||:::||||||
Db   1122 GCATATGTGATCTTTGATATTATTGATAATAATCCTAAAATTGACAGTTTTTCAGAGAGA 1181

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
          |||:::||||||:::||||||:::|||   :::||||||   :::|||   |||:::|||
Db   1182 GGACACAAACCAGACAGCATCAAAGGGAATTTGGAGTTCAATGATGTTCACTTTTCTTAC 1241

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
          |||||||||   ::::::||||||||||||||||||||||||:::::::::||||||:::
Db   1242 CCTTCTCGAGCTAACGTCAAGATCTTGAAGGGCCTCAACCTGAAGGTGCAGAGTGGGCAG 1301

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
          ||||||||||||||||||   :::|||   ||||||||||||   |||||||||:::|||
Db   1302 ACGGTGGCCCTGGTTGGAAGTAGTGGCTGTGGGAAGAGCACAACGGTCCAGCTGATACAG 1361

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
          ||||||||||||||||||:::|||   |||   :::|||      |||||||||
Db   1362 AGGCTCTATGACCCTGATGAGGGCACAATTAACATTGATGGGCAGGATATTAGGAACTTT 1421

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
          ||||||   :::   |||:::   |||||||||||||||||||||||||||||||
Db   1422 AATGTAAACTATCTGAGGGAAATCATTGGTGTGGTGAGTCAGGAGCCGGTGCTGTTTTCC 1481

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
          |||||||||:::   ||||||   |||||||||   :::||||||   :::|||::::::
Db   1482 ACCACAATTGCTGAAAATATTTGTTATGGCCGTGGAAATGTAACCATGGATGAGATAAAG 1541

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
          :::|||   :::|||||||||||||||:::|||||||||:::      |||   ||||||:::
Db   1542 AAAGCTGTCAAAGAGGCCAACGCCTATGAGTTTATCATGAAATTACCACAGAAATTTGAC 1601

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
          ||||||||||||:::||||||:::||||||||||||||||||||||||||||||||||
Db   1602 ACCCTGGTTGGAGAGAGAGGGGCCCAGCTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 1661

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
          ||||||||||||||||||||||||:::||||||||||||||||||||||||||||||||
Db   1662 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTTCTGCTGGATGAGGCCACGTCAGCATTG 1721

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
          |||:::|||||||||::::::   |||||||||||||||:::||||||   :::|||||||||
Db   1722 GACACAGAAAGTGAAGCTGAGGTACAGGCAGCTCTGGATAAGGCCAGAGAAGGCCGGACC 1781

Qy    541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
          |||||||||:::||||||||||||||||||:::|||:::||||||:::|||
Db   1782 ACCATTGTGATAGCACACCGACTGTCTACGGTCCGAAATGCAGATGTCATCGCTGGGTTT 1841

Qy    561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
          :::||||||::::::   |||:::|||:::|||:::|||||||||   |||   |||:::
Db   1842 GAGGATGGAGTAATTGTGGAGCAAGGAAGCCACAGCGAACTGATGAAGAAGGAAGGGGTG 1901

Qy    581 TyrTyrSerLeuVal--------------------------MetSerGlnAspIleLys 591
          |||:::   ||||||                          |||::::::   :::
Db   1902 TACTTCAAACTTGTCAACATGCAGACATCAGGAAGCCAGATCCAGTCAGAAGAATTTGAA 1961

Qy    592 LysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu 611
                |||||||:::      :::                     |||:::
Db   1962 CTAAATGATGAAAAGGCTGCCACTAGAATGGCCCCAAATGGCTGGAAATCTCGCCTATTT 2021
```

```
Qy    612 ProLeuHisSerValLysSerIleLys---SerAspPheIleAspLysAla---GluGlu 629
Db   2022 AGGCATTCTACTCAGAAAAACCTTAAAAATTCACAAATGTGTCAGAAGAGCCTTGATGTG 2081

Qy    630 SerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeu 649
Db   2082 GAAACCGATGGACTTGAAGCAAATGTGCCACCAGTGTCCTTTCTGAAGGTCCTGAAACTG 2141

Qy    650 AsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThr 669
Db   2142 AATAAAACAGAATGGCCCTACTTTGTCGTGGGAACAGTATGTGCCATTGCCAATGGGGGG 2201

Qy    670 ValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAsp 689
Db   2202 CTTCAGCCGGCATTTTCAGTCATATTCTCAGAGATCATAGCGATTTTTGGACCAGGCGAT 2261

Qy    690 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 709
Db   2262 GATGCAGTGAAGCAGCAGAAGTGCAACATATTCTCTTTGATTTTCTTATTTCTGGGAATT 2321

Qy    710 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 729
Db   2322 ATTTCTTTTTTACTTTCTTCCTTCAGGGTTTCACGTTTGGGAAAGCTGGCGAGATCCTC 2381

Qy    730 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 749
Db   2382 ACCAGAAGACTGCGGTCAATGGCTTTTAAAGCAATGCTAAGACAGGACATGAGCTGGTTT 2441

Qy    750 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 769
Db   2442 GATGACCATAAAAACAGTACTGGTGCACTTTCTACAAGACTTGCCACAGATGCTGCCCAA 2501

Qy    770 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 789
Db   2502 GTCCAAGGAGCCACAGGAACCAGGTTGGCTTTAATTGCACAGAATATAGCTAACCTTGGA 2561

Qy    790 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 809
Db   2562 ACTGGTATTATCATATCATTTATCTACGGTTGGCAGTTAACCCTATTGCTATTAGCAGTT 2621

Qy    810 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 829
Db   2622 GTTCCAATTATTGCTGTGTCAGGAATTGTTGAAATGAAATTGTTGGCTGGAAATGCCAAA 2681

Qy    830 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 849
Db   2682 AGAGATAAAAAAGAACTGGAAGCTGCTGGAAAGATTGCAACAGAGGCAATAGAAAATATT 2741

Qy    850 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 869
Db   2742 AGGACAGTTGTGTCTTTGACCCAGGAAAGAAAATTTGAATCAATGTATGTTGAAAAAATTG 2801

Qy    870 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 889
Db   2802 TATGGACCTTACAGGAATTCTGTGCAGAAGGCACACATCTATGGAATTACTTTTAGTATC 2861

Qy    890 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 909
Db   2862 TCACAAGCATTTATGTATTTTTCCTATGCCGGTTGTTTTCGATTTGGTGCATATCTCATT 2921

Qy    910 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 929
Db   2922 GTGAATGGACATATGCGCTTCAGAGATGTTATTCTGGTGTTTTCTGCAATTGTATTTGGT 2981

Qy    930 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 949
Db   2982 GCAGTGGCTCTAGGACATGCCAGTTCATTTGCTCCAGACTATGCTAAAGCTAAGCTGTCT 3041

Qy    950 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 969
Db   3042 GCAGCCCACTTATTCATGCTGTTTGAAAGACAACCTCTGATTGACAGCTACAGTGAAGAG 3101

Qy    970 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 989
Db   3102 GGGCTGAAGCCTGATAAATTTGAAGGAAATATAACATTTAATGAAGTCGTGTTCAACTAT 3161

Qy    990 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 1009
Db   3162 CCCACCCGAGCAAACGTGCCAGTGCTTCAGGGGCTGAGCCTGGAGGTGAAGAAAGGCCAG 3221

Qy   1010 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 1029
Db   3222 ACACTAGCCCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACGGTGGTCCAGCTCCTGGAG 3281

Qy   1030 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 1049
Db   3282 CGGTTCTACGACCCCTTGGCGGGGACAGTGCTTCTCGATGGTCAAGAAGCAAAGAAACTC 3341

Qy   1050 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 1069
Db   3342 AATGTCCAGTGGCTCAGAGCTCAACTCGGAATCGTGTCTCAGGAGCCTATCCTATTTGAC 3401

Qy   1070 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 1089
Db   3402 TGCAGCATTGCCGAGAATATTGCCTATGGAGACAACAGCCGGGTTGTATCACAGGATGAA 3461

Qy   1090 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 1109
Db   3462 ATTGTGAGTGCAGCCAAAGCTGCCAACATACATCCTTTCATCGAGACGTTACCCCACAAA 3521

Qy   1110 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1129
Db   3522 TATGAAACAAGAGTGGGAGATAAGGGGACTCAGCTCTCAGGAGGTCAAAAACAGAGGATT 3581

Qy   1130 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 1149
Db   3582 GCTATTGCCCGAGCCCTCATCAGACAACCTCAAATCCTCCTGTTGGATGAAGCTACATCA 3641

Qy   1150 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1169
Db   3642 GCTCTGGATACTGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 3701

Qy   1170 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1189
Db   3702 CGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTG 3761

Qy   1190 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1209
Db   3762 GTGTTTCAGAATGGGAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAA 3821

Qy   1210 AspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db   3822 GGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3854
```

Search completed: March 31, 2003, 16:09:09
Job time : 711.798 secs

```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 31, 2003, 02:24:20 ; Search time 178.956 Seconds
                                           (without alignments)
                                           2094.145 Million cell updates/sec

Title:          US-09-873-409-5
Perfect score:  6159
Sequence:       1 MILGILASLVNGACLPLMPL..........QELLRNRDIYFKLVNAQSVQ 1222

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters:     882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115422_19259/app_query.fasta_1.7544
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_578_@runat_27032003_115422_19259 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Issued_Patents_NA:*
                1:  /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
                2:  /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
                3:  /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
                4:  /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
                5:  /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
                6:  /cgn2_6/ptodata/2/ina/backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3521.5	57.2	4646	1	US-08-181-471-2	Sequence 2, Appli
2	3521.5	57.2	6505	2	US-08-793-610-5	Sequence 5, Appli
3	3521.5	57.2	9318	2	US-08-793-610-6	Sequence 6, Appli
4	3520	57.2	4264	2	US-08-784-649A-1	Sequence 1, Appli
5	3520	57.2	4264	2	US-08-784-649A-5	Sequence 5, Appli
6	3519.5	57.1	4669	6	5206352-3	Patent No. 5206352
7	3503.5	56.9	4669	2	US-08-752-447-1	Sequence 1, Appli
8	3503.5	56.9	4669	4	US-09-316-167-1	Sequence 1, Appli
9	3494	56.7	4233	3	US-09-120-513-1	Sequence 1, Appli
10	3494	56.7	4233	4	US-09-450-105-1	Sequence 1, Appli
11	3473.5	56.4	4669	2	US-08-583-276-18	Sequence 18, Appl
12	2280	37.0	2726	1	US-08-461-823-1	Sequence 1, Appli
13	2259	36.7	4002	2	US-08-996-545-1	Sequence 1, Appli
14	2259	36.7	4002	2	US-08-996-545-3	Sequence 3, Appli
15	2259	36.7	4002	4	US-09-328-320-1	Sequence 1, Appli
16	2259	36.7	4002	4	US-09-328-320-3	Sequence 3, Appli
17	2250	36.5	4047	2	US-08-612-734B-1	Sequence 1, Appli
18	2187.5	35.5	4800	2	US-08-612-734B-3	Sequence 3, Appli
19	2143.5	34.8	3924	1	US-08-395-246C-1	Sequence 1, Appli
20	2100.5	34.1	4224	1	US-08-612-521-1	Sequence 1, Appli
21	1722	28.0	6143	1	US-08-612-521-3	Sequence 3, Appli
22	1680.5	27.3	3924	2	US-08-996-644-3	Sequence 3, Appli
23	1680.5	27.3	3924	3	US-09-352-552-3	Sequence 3, Appli
24	1680.5	27.3	3927	2	US-08-996-644-1	Sequence 1, Appli
25	1680.5	27.3	3927	3	US-09-352-552-1	Sequence 1, Appli
26	1646.5	26.7	3909	1	US-08-232-537-1	Sequence 1, Appli
27	1636	26.6	3792	4	US-09-351-224E-10	Sequence 10, Appl
28	1501.5	24.4	3999	4	US-09-351-224E-9	Sequence 9, Appli
29	1106.5	18.0	13188	4	US-08-961-527-70	Sequence 70, Appl
30	1047.5	17.0	7760	4	US-08-961-527-63	Sequence 63, Appl
31	989	16.1	4411529	4	US-09-103-840A-1	Sequence 1, Appli
32	987	16.0	4403765	4	US-09-103-840A-2	Sequence 2, Appli
33	940	15.3	6492	4	US-08-961-527-188	Sequence 188, App
34	938	15.2	7186	4	US-08-961-527-39	Sequence 39, Appl
35	738	12.0	2376	1	US-08-394-880B-1	Sequence 1, Appli
36	708.5	11.5	1749	4	US-09-134-001C-1893	Sequence 1893, Ap
37	695	11.3	5889	1	US-08-463-092B-5	Sequence 5, Appli
38	695	11.3	5889	2	US-08-462-109A-5	Sequence 5, Appli
39	695	11.3	5889	2	US-08-460-907B-5	Sequence 5, Appli
40	695	11.3	5889	3	US-08-463-179A-5	Sequence 5, Appli
41	695	11.3	5889	3	US-08-461-384B-5	Sequence 5, Appli
42	687	11.2	5011	1	US-08-463-092B-3	Sequence 3, Appli
43	687	11.2	5011	2	US-08-462-109A-3	Sequence 3, Appli
44	687	11.2	5011	2	US-08-460-907B-3	Sequence 3, Appli
45	687	11.2	5011	3	US-08-463-179A-3	Sequence 3, Appli

ALIGNMENTS

```
RESULT 1
US-08-181-471-2
; Sequence 2, Application US/08181471
; Patent No. 5641508
;  GENERAL INFORMATION:
;    APPLICANT:  Li, Lingna
;    APPLICANT:  Lishko, Valeryi K.
;    TITLE OF INVENTION:  METHOD FOR DELIVERING BENEFICIAL
;    TITLE OF INVENTION:  COMPOSITIONS TO HAIR FOLLICLES
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Thomas Fitting
;      STREET:  12526 High Bluff Drive, Suite 300
;      CITY:  San Diego
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92130
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/181,471
;      FILING DATE:  13-JAN-1994
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/041,553
;      FILING DATE:  02-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitting, Thomas
;      REGISTRATION NUMBER:  34,163
;      REFERENCE/DOCKET NUMBER:  ANT0029P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-792-3680
;      TELEFAX:  619-792-8477
```

```
;  INFORMATION FOR SEQ ID NO:  2:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  4646 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;   MOLECULE TYPE:  cDNA
;   HYPOTHETICAL:  NO
;   ANTI-SENSE:  NO
;   FEATURE:
;     NAME/KEY:  CDS
;     LOCATION:  425..4267
US-08-181-471-2

Alignment Scores:
Pred. No.:               0          Length:        4646
Score:                   3521.50    Matches:       681
Percent Similarity:      73.92%     Conservative:  246
Best Local Similarity:   54.31%     Mismatches:    264
Query Match:             57.18%     Indels:        63
DB:                      1          Gaps:          10

US-09-873-409-5 (1-1222) x US-08-181-471-2 (1-4646)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
             |||::::::|||  ||||||::::::::::::||||||  ||||||||||||   |||
Db    575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy     21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
          |||   |||||||||:::|||                          :::|||:::|||
Db    635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                   ::::::        |||        ||||||
Db    695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                |||       ||||||  ||||||||     |||:::             |||
Db    755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
          ||||||:::|||  |||   :::   |||   ||||||        :::||||||||||||
Db    815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
          |||||||||:::::::::   |||:::||||||||||||||        |||:::||||||
Db    875 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
          ||||||||||||:::|||   |||:::   |||||||::::::   ||||||||||||||
Db    935 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 994

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
                :::    ||||||:::|||:::||||||         |||      ||||||      :::
Db    995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
          ||||||||||||||||||||||        |||:::   ||||||:::::::::  ||||||||
Db   1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
                ::::::::::::   |||   |||   |||||||||   ||||||:::||||||
Db   1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
          |||||||||||||||||||||||:::::::||||||||||||||||||||||      |||
Db   1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
          :::||||||||||:::
Db   1235 AAGAAAGAACTTGAA--------------------------------------------- 1249

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                                      ||||||    :::||||||::::::
Db   1250 -----------------------------------AGGTACAACAAAAATTTAGAAGAA 1273

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
          ||||||   |||||||||:::   |||   :::   :::|||:::||||||   :::
Db   1274 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 1333

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
             :::       :::|||   ||||||||||||||||||||||:::|||:::|||:::
Db   1334 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 1393

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
          ||||||      |||:::||||||   ||||||   ||||||||||||||:::
Db   1394 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1447

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
          ::::::   :::|||   |||   |||      |||   ||||||   ||||||||||||
Db   1448 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1507

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
          |||:::   ||||||::::::||||||   ||||||||||||||::::::|||   :::
Db   1508 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1567

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
          |||:::|||||||||::::::|||:::|||   :::||||||:::||||||   |||:::|||
Db   1568 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1627

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
          ||||||||   ::::|||||||||||||||||||||||||::::::::::||||||:::
Db   1628 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1687

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
          ||||||||||||||||     :::|||   ||||||||||||    |||||||||:::|||
Db   1688 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1747

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
          ||||||||||||||     :::|||   :::   ||||||        |||||||||   :::
Db   1748 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1807

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
          |||||||||     ||||:::   ||||||||||||||||||||||||||||||||||
Db   1808 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1867

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
          |||||||||:::   ||||||:::||||||||||||:::|||||||     :::|||:::|||
Db   1868 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1927

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
          :::|||   :::|||||||||||||||||||||||||||:::      |||:::||||||:::
Db   1928 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1987

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
          |||||||||||||||:::||||||||||:::|||||||||||||||||||||||||||||
Db   1988 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 2047

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
          ||||||||||||||||||||||||||||||:::||||||||||||||||||||||||
Db   2048 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 2107

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
          |||:::||||||::::::   ||||||   ||||||:::||||||   ||||||||||||
Db   2108 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACC 2167

Qy    541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
          |||||||||:::||||||||||||||||||:::|||:::||||||:::|||
Db   2168 ACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC 2227

Qy    561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
             ||||||::::::   |||||||||   |||   |||||||||   ::::::|||:::
Db   2228 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATT 2287

Qy    581 TyrTyrSerLeuValMetSerGln-----------------AspIleLysLysAlaAsp 594
          |||:::   ||||||   |||                 :::::::::   |||
```

```
Db   2288 TACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 2347

Qy    595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
Db   2348 GATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGT 2407

Qy    615 SerVal------------------LysSerIleLys---SerAspPheIleAspLysAla 627
Db   2408 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAG 2467

Qy    628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
Db   2468 CTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATG 2527

Qy    648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
Db   2528 AAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAAT 2587

Qy    668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
Db   2588 GGAGGCCTGCAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGA 2647

Qy    688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db   2648 ATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCC 2707

Qy    707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db   2708 CTTGGAATTATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGA 2767

Qy    727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db   2768 GAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTG 2827

Qy    747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
Db   2828 AGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGAT 2887

Qy    767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db   2888 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCA 2947

Qy    787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
Db   2948 AATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTC 3007

Qy    807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db   3008 TTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGA 3067

Qy    827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db   3068 CAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 3127

Qy    847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db   3128 GAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCT 3187

Qy    867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db   3188 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 3247

Qy    887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db   3248 TTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCC 3307

Qy    907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db   3308 TACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTT 3367

Qy    927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db   3368 GTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCC 3427

Qy    947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db   3428 AAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTAC 3487

Qy    967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db   3488 AGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTA 3547

Qy    987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db   3548 TTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAG 3607

Qy   1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
Db   3608 AAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAG 3667

Qy   1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
Db   3668 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATA 3727

Qy   1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
Db   3728 AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATC 3787

Qy   1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db   3788 CTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCA 3847

Qy   1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db   3848 CAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTG 3907

Qy   1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
Db   3908 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAA 3967

Qy   1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
Db   3968 CAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAA 4027

Qy   1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
Db   4028 GCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCC 4087

Qy   1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db   4088 AGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGAC 4147

Qy   1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
Db   4148 TTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTG 4207

Qy   1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db   4208 GCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 4249
```

RESULT 2
US-08-793-610-5
; Sequence 5, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
; APPLICANT: BAUM, Christopher
; APPLICANT: STOCKING-HARBERS, Carol
; APPLICANT: OSTERTAG, Wolfram
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEROF
; TITLE OF INVENTION: FOR GENE TRANSFER
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.

```
;      COUNTRY:  U.S.A.
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/793,610
;      FILING DATE:  07-MAR-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P 44 31 973.8
;      FILING DATE:  08-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 195 03 952.1
;      FILING DATE:  07-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/03175
;      FILING DATE:  10-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berman, Richard J.
;      REGISTRATION NUMBER:  39,105
;      REFERENCE/DOCKET NUMBER:  P1614-7007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)638-5000
;      TELEFAX:  (202)638-4810
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6505 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA
US-08-793-610-5

Alignment Scores:
Pred. No.:              0          Length:         6505
Score:                  3521.50    Matches:        681
Percent Similarity:     73.92%     Conservative:   246
Best Local Similarity:  54.31%     Mismatches:     264
Query Match:            57.18%     Indels:         63
DB:                     2          Gaps:           10

US-09-873-409-5 (1-1222) x US-08-793-610-5 (1-6505)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db   1967 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 2026

Qy     21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db   2027 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 2086

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db   2087 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 2146

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db   2147 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 2206

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db   2207 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 2266

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db   2267 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 2326

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db   2327 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 2386

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db   2387 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 2446

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db   2447 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 2506

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db   2507 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 2566

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db   2567 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 2626

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db   2627 AAGAAAGAACTTGAA--------------------------------------------- 2641

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db   2642 ------------------------------------AGGTACAACAAAAATTTAGAAGAA 2665

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db   2666 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 2725

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db   2726 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 2785

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db   2786 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 2839

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db   2840 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 2899

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   2900 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 2959

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   2960 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 3019

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   3020 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 3079

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   3080 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 3139

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   3140 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 3199

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   3200 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 3259

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   3260 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 3319

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   3320 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 3379

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
```

```
Db   3380 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 3439

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   3440 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 3499

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db   3500 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACC 3559

Qy    541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db   3560 ACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC 3619

Qy    561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db   3620 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATT 3679

Qy    581 TyrTyrSerLeuValMetSerGln------------------AspIleLysLysAlaAsp 594
Db   3680 TACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 3739

Qy    595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
Db   3740 GATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGT 3799

Qy    615 SerVal------------------LysSerIleLys---SerAspPheIleAspLysAla 627
Db   3800 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAG 3859

Qy    628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
Db   3860 CTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATG 3919

Qy    648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
Db   3920 AAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAAT 3979

Qy    668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
Db   3980 GGAGGCCTGCAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGA 4039

Qy    688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db   4040 ATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCC 4099

Qy    707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db   4100 CTTGGAATTATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGA 4159

Qy    727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db   4160 GAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTG 4219

Qy    747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
Db   4220 AGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGAT 4279

Qy    767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db   4280 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCA 4339

Qy    787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
Db   4340 AATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTC 4399

Qy    807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db   4400 TTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGA 4459

Qy    827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db   4460 CAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 4519

Qy    847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db   4520 GAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCT 4579

Qy    867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db   4580 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 4639

Qy    887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db   4640 TTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCC 4699

Qy    907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db   4700 TACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTT 4759

Qy    927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db   4760 GTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCC 4819

Qy    947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db   4820 AAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTAC 4879

Qy    967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db   4880 AGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTA 4939

Qy    987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db   4940 TTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAG 4999

Qy   1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
Db   5000 AAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAG 5059

Qy   1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
Db   5060 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATA 5119

Qy   1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
Db   5120 AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATC 5179

Qy   1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db   5180 CTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCA 5239

Qy   1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db   5240 CAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTG 5299

Qy   1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
Db   5300 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAA 5359

Qy   1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
Db   5360 CAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAA 5419

Qy   1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
Db   5420 GCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCC 5479

Qy   1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db   5480 AGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGAC 5539

Qy   1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
Db   5540 TTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTG 5599
```

```
Qy   1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
                :::    |||||||||   :::|||:::   |||:::
Db   5600 GCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 5641

RESULT 3
US-08-793-610-6
; Sequence 6, Application US/08793610
; Patent No. 5858744
;  GENERAL INFORMATION:
;    APPLICANT:  BAUM, Christopher
;    APPLICANT:  STOCKING-HARBERS, Carol
;    APPLICANT:  OSTERTAG, Wolfram
;    TITLE OF INVENTION:  RETROVIRAL VECTOR HYBRIDS AND THE USE THEROF
;    TITLE OF INVENTION:  FOR GENE TRANSFER
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nikaido, Marmelstein, Murray & Oram LLP
;      STREET:  655 Fifteenth Street N.W. Suite 330
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/793,610
;      FILING DATE:  07-MAR-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P 44 31 973.8
;      FILING DATE:  08-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 195 03 952.1
;      FILING DATE:  07-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/03175
;      FILING DATE:  10-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berman, Richard J.
;      REGISTRATION NUMBER:  39,105
;      REFERENCE/DOCKET NUMBER:  P1614-7007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)638-5000
;      TELEFAX:  (202)638-4810
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9318 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA
US-08-793-610-6

Alignment Scores:
Pred. No.:               0          Length:          9318
Score:                   3521.50    Matches:         681
Percent Similarity:      73.92%     Conservative:    246
Best Local Similarity:   54.31%     Mismatches:      264
Query Match:             57.18%     Indels:          63
DB:                      2          Gaps:            10

US-09-873-409-5 (1-1222) x US-08-793-610-6 (1-9318)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          |||::::::|||   |||||::::::::::::||||||   |||||||||||   |||
Db   1926 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 1985

Qy     21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
          |||   |||||||||:::|||                           :::|||:::|||
Db   1986 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 2045

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                   ::::::         |||         ||||||
Db   2046 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 2105

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                |||          |||||   ||||||||         |||:::        |||
Db   2106 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 2165

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
          ||||||:::|||   |||   :::   |||   ||||||      :::|||||||||||||
Db   2166 ATTCAGGTTTCATTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 2225

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
          ||||||||:::::::::   |||:::|||||||||||||||      |||:::||||||
Db   2226 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 2285

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
          ||||||||||||:::|||   |||:::   ||||||::::::   ||||||||||||||
Db   2286 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 2345

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
             :::   ||||||:::|||:::||||||      |||      ||||||      :::
Db   2346 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 2405

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
          ||||||||||||||||||||   |||:::   ||||||::::::     |||||||||
Db   2406 GGTTGGAAGCTAACCCTTGTGATTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 2465

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
             ::::::::::::   |||   |||   |||||||||   ||||||:::||||||
Db   2466 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 2525

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
          |||||||||||||||||||||||||:::::||||||||||||||||||||      |||
Db   2526 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 2585

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
          :::|||||||||:::
Db   2586 AAGAAAGAACTTGAA--------------------------------------------- 2600

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                               ||||||   :::||||||::::::
Db   2601 ------------------------------------AGGTACAACAAAAATTTAGAAGAA 2624

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
          ||||||   |||||||||:::   |||   :::   :::|||:::||||||      :::
Db   2625 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 2684

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
             :::      :::|||   ||||||||||||||||||||||:::|||:::|||:::
Db   2685 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 2744

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
          ||||||      |||:::||||||   ||||||   |||||||||||||||:::
Db   2745 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 2798

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
          ::::::   :::|||   |||   |||      |||   ||||||   ||||||||||||
Db   2799 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 2858

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
          |||:::   ||||||::::::||||||   |||||||||||||||::::::|||   :::
Db   2859 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 2918

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
          |||:::||||||::::::|||:::|||   :::||||||:::||||||   |||:::|||
Db   2919 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 2978

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
          ||||||||      :::||||||||||||||||||||||||::::::::::||||||:::
Db   2979 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 3038
```

```
Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   3039 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 3098

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   3099 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 3158

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   3159 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 3218

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   3219 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 3278

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   3279 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 3338

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   3339 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 3398

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   3399 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 3458

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db   3459 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACC 3518

Qy    541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db   3519 ACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC 3578

Qy    561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db   3579 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATT 3638

Qy    581 TyrTyrSerLeuValMetSerGln------------------AspIleLysLysAlaAsp 594
Db   3639 TACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 3698

Qy    595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
Db   3699 GATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGT 3758

Qy    615 SerVal------------------LysSerIleLys---SerAspPheIleAspLysAla 627
Db   3759 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAG 3818

Qy    628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
Db   3819 CTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATG 3878

Qy    648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
Db   3879 AAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAAT 3938

Qy    668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
Db   3939 GGAGGCCTGCAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGA 3998

Qy    688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db   3999 ATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCC 4058

Qy    707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db   4059 CTTGGAATTATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGA 4118

Qy    727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db   4119 GAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTG 4178

Qy    747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
Db   4179 AGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGAT 4238

Qy    767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db   4239 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCA 4298

Qy    787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
Db   4299 AATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTC 4358

Qy    807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db   4359 TTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGA 4418

Qy    827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db   4419 CAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 4478

Qy    847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db   4479 GAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCT 4538

Qy    867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db   4539 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 4598

Qy    887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db   4599 TTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCC 4658

Qy    907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db   4659 TACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTT 4718

Qy    927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db   4719 GTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCC 4778

Qy    947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db   4779 AAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTAC 4838

Qy    967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db   4839 AGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTA 4898

Qy    987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db   4899 TTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAG 4958

Qy   1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
Db   4959 AAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAG 5018

Qy   1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
Db   5019 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATA 5078

Qy   1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
Db   5079 AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATC 5138

Qy   1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db   5139 CTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCA 5198

Qy   1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
```

```
Db     5199 CAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTG 5258

Qy     1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
            |||   |||||:::||:::|||||   |||||    ||||||||||||||||||
Db     5259 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAA 5318

Qy     1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
            ||||||:::||||||||||||||||||::::::::|||   ||||||||||||||||
Db     5319 CAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTGCTTTTGGATGAA 5378

Qy     1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
            ||||||||||||||||||   :::||||||||||||||||||   ||||||||||||||
Db     5379 GCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCC 5438

Qy     1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
            |||   ||||||||||||:::|||:::   ||||||||||||   ||||||||||||||
Db     5439 AGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGAC 5498

Qy     1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
            ||||||||||||      ||||||::::::||||||   ||||||||||||:::||||||
Db     5499 TTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTG 5558

Qy     1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
               :::   |||||||||   :::|||:::   |||:::
Db     5559 GCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 5600
```

RESULT 4
US-08-784-649A-1
; Sequence 1, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
;   APPLICANT: Sikic, Branimir I
;   APPLICANT: Chen, Gang
;   TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
;   TITLE OF INVENTION: CYCLOSPORIN MODULATION
;   NUMBER OF SEQUENCES: 5
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Fish & Richardson
;   STREET: 2200 Sand Hill Road
;   CITY: Menlo Park
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94025
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/784,649A
;   FILING DATE:
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Sherwood, Pamela J
;   REGISTRATION NUMBER: Reg.No. 5830697 36,677
;   REFERENCE/DOCKET NUMBER: 06037/007001
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415-322-5070
;   TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4264 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
US-08-784-649A-1

```
Alignment Scores:
Pred. No.:               0              Length:         4264
Score:                   3520.00        Matches:        682
Percent Similarity:      74.00%         Conservative:   246
Best Local Similarity:   54.39%         Mismatches:     262
Query Match:             57.15%         Indels:         64
DB:                      2              Gaps:           11

US-09-873-409-5 (1-1222) x US-08-784-649A-1 (1-4264)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            |||::::::|||   ||||||::::::::::::||||||   |||||||||||   |||
Db      292 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 351

Qy       21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
            |||   |||||||||:::|||                           :::|||:::|||
Db      352 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 411

Qy       32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                        ::::::      |||         ||||||
Db      412 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 471

Qy       44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                  |||         ||||||   |||||||||         |||:::      |||
Db      472 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 531

Qy       62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
            ||||||:::|||   |||   :::   |||   ||||||         :::||||||||||||
Db      532 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 591

Qy       82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
            ||||||||:::::::::   |||:::||||||||||||||||   |||:::||||||
Db      592 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 651

Qy      102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
            ||||||||||||:::|||   |||:::   ||||||::::::||||||||||||||||||
Db      652 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAAATT 711

Qy      121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
               :::||||||:::|||:::||||||         |||      ||||||      :::
Db      712 GGA---ATGTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 768

Qy      141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
            ||||||||||||||||||||   |||:::   ||||||::::::      |||||||||
Db      769 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 828

Qy      161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
               ::::::::::::   |||   |||   |||||||||   ||||||:::||||||
Db      829 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 888

Qy      181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
            ||||||||||||||||||||||||::::::|||||||||||||||||||      |||
Db      889 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 948

Qy      201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
            :::|||||||||:::
Db      949 AAGAAAGAACTTGAA--------------------------------------------- 963

Qy      221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                               ||||||   :::||||||::::::
Db      964 -----------------------------------AGGTACAACAAAAATTTAGAAGAA 987

Qy      241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
            ||||||   |||||||||:::   |||   :::   :::|||:::||||||   :::
Db      988 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 1047

Qy      261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
               :::      :::|||   ||||||||||||||||||||||:::|||:::|||:::
Db     1048 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 1107

Qy      281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
            ||||||      |||:::||||||   ||||||   |||||||||||||||:::
Db     1108 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1161

Qy      301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
            ::::::   :::|||   |||   |||         |||   ||||||   ||||||||||||
Db     1162 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1221
```

```
Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   1222 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1281

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   1282 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1341

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   1342 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1401

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   1402 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1461

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   1462 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1521

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   1522 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1581

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   1582 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1641

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   1642 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1701

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   1702 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 1761

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   1762 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 1821

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db   1822 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACC 1881

Qy    541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db   1882 ACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC 1941

Qy    561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db   1942 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATT 2001

Qy    581 TyrTyrSerLeuValMetSerGln------------------AspIleLysLysAlaAsp 594
Db   2002 TACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 2061

Qy    595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
Db   2062 GATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGT 2121

Qy    615 SerVal------------------LysSerIleLys---SerAspPheIleAspLysAla 627
Db   2122 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAG 2181

Qy    628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
Db   2182 CTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATG 2241

Qy    648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
Db   2242 AAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAAT 2301

Qy    668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
Db   2302 GGAGGCCTGCAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGA 2361

Qy    688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db   2362 ATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCC 2421

Qy    707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db   2422 CTTGGAATTATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGA 2481

Qy    727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db   2482 GAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTG 2541

Qy    747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
Db   2542 AGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGAT 2601

Qy    767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db   2602 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCA 2661

Qy    787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
Db   2662 AATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTC 2721

Qy    807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db   2722 TTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGA 2781

Qy    827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db   2782 CAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 2841

Qy    847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db   2842 GAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCT 2901

Qy    867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db   2902 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 2961

Qy    887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db   2962 TTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCC 3021

Qy    907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db   3022 TACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTT 3081

Qy    927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db   3082 GTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCC 3141

Qy    947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db   3142 AAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTAC 3201

Qy    967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db   3202 AGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTA 3261

Qy    987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db   3262 TTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAG 3321

Qy   1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
Db   3322 AAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAG 3381

Qy   1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
```

```
Db    3382 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATA 3441

Qy    1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
Db    3442 AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATC 3501

Qy    1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db    3502 CTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCA 3561

Qy    1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db    3562 CAGGAAGAGATTGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTG 3621

Qy    1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
Db    3622 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAA 3681

Qy    1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
Db    3682 CAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAA 3741

Qy    1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
Db    3742 GCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCC 3801

Qy    1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db    3802 AGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGAC 3861

Qy    1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
Db    3862 TTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTG 3921

Qy    1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db    3922 GCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3963
```

RESULT 5
US-08-784-649A-5
; Sequence 5, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
;   APPLICANT: Sikic, Branimir I
;   APPLICANT: Chen, Gang
;   TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
;   TITLE OF INVENTION: CYCLOSPORIN MODULATION
;   NUMBER OF SEQUENCES: 5
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Fish & Richardson
;   STREET: 2200 Sand Hill Road
;   CITY: Menlo Park
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94025
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/784,649A
;   FILING DATE:
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Sherwood, Pamela J
;   REGISTRATION NUMBER: Reg.No. 5830697 36,677
;   REFERENCE/DOCKET NUMBER: 06037/007001
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415-322-5070
;   TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 5:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4264 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
US-08-784-649A-5

Alignment Scores:

Pred. No.:	0	Length:	4264
Score:	3520.00	Matches:	682
Percent Similarity:	74.00%	Conservative:	246
Best Local Similarity:	54.39%	Mismatches:	262
Query Match:	57.15%	Indels:	64
DB:	2	Gaps:	11

US-09-873-409-5 (1-1222) x US-08-784-649A-5 (1-4264)

```
Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db     292 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 351

Qy      21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db     352 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 411

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db     412 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 471

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db     472 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 531

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db     532 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 591

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db     592 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 651

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db     652 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAAATT 711

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db     712 GGA---ATGTTCCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 768

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db     769 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 828

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db     829 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 888

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db     889 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 948

Qy     201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db     949 AAGAAAGAACTTGAA--------------------------------------------- 963

Qy     221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db     964 ------------------------------------AGGTACAACAAAAATTTAGAAGAA 987

Qy     241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db     988 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 1047
```

```
Qy     261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db    1048 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 1107

Qy     281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db    1108 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1161

Qy     301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db    1162 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1221

Qy     321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db    1222 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1281

Qy     341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db    1282 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1341

Qy     361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db    1342 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1401

Qy     381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db    1402 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1461

Qy     401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db    1462 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1521

Qy     421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db    1522 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1581

Qy     441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db    1582 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1641

Qy     461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db    1642 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1701

Qy     481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db    1702 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 1761

Qy     501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db    1762 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 1821

Qy     521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db    1822 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACC 1881

Qy     541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db    1882 ACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC 1941

Qy     561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db    1942 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATT 2001

Qy     581 TyrTyrSerLeuValMetSerGln------------------AspIleLysLysAlaAsp 594
Db    2002 TACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 2061

Qy     595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
Db    2062 GATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGT 2121

Qy     615 SerVal------------------LysSerIleLys---SerAspPheIleAspLysAla 627
Db    2122 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAG 2181

Qy     628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
Db    2182 CTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATG 2241

Qy     648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
Db    2242 AAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAAT 2301

Qy     668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
Db    2302 GGAGGCCTGCAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGA 2361

Qy     688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db    2362 ATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCC 2421

Qy     707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db    2422 CTTGGAATTATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGA 2481

Qy     727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db    2482 GAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTG 2541

Qy     747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
Db    2542 AGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGAT 2601

Qy     767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db    2602 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCA 2661

Qy     787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
Db    2662 AATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTC 2721

Qy     807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db    2722 TTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGA 2781

Qy     827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db    2782 CAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 2841

Qy     847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db    2842 GAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCT 2901

Qy     867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db    2902 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 2961

Qy     887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db    2962 TTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCC 3021

Qy     907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db    3022 TACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTT 3081

Qy     927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db    3082 GTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCC 3141

Qy     947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db    3142 AAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTAC 3201

Qy     967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
```

```
Db    3202 AGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTA 3261

Qy     987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
           |||   |||||   ||||||||:::   :::|||:::||||||||||   ::::::
Db    3262 TTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAG 3321

Qy    1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
           :::|||:::|||:::|||   ||||||||||||||||||||||||||||   ||||||
Db    3322 AAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAG 3381

Qy    1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
           ||||||:::|||   |||||||||:::   |||:::||||||   ||||||   :::
Db    3382 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATA 3441

Qy    1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
           |||   |||||||||||||||||||||:::   :::   ||||||   |||||||||:::
Db    3442 AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATC 3501

Qy    1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
           ||||||:::||||||||||||||||||||||||||||||||||||||||||||||
Db    3502 CTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCA 3561

Qy    1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
              :::||||||         ||||||      |||||||||||:::|||||||||   |||
Db    3562 CAGGAAGAGATTGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTG 3621

Qy    1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
           |||   ||||||:::|||:::||||||   ||||||   ||||||||||||||||||||||
Db    3622 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAA 3681

Qy    1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
           ||||||:::||||||||||||||||||||:::::::::|||   ||||||||||||||||||
Db    3682 CAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAA 3741

Qy    1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
           ||||||||||||||||||   :::||||||||||||||||||   |||||||||||||||
Db    3742 GCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCC 3801

Qy    1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
           |||   ||||||||||||:::|||:::   ||||||||||||   ||||||||||||||||
Db    3802 AGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGAC 3861

Qy    1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
           ||||||||||||         ||||||::::::||||||   ||||||||||||:::||||||
Db    3862 TTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTG 3921

Qy    1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
              :::   ||||||||   :::|||:::   |||:::
Db    3922 GCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3963
```

RESULT 6
5206352-3
;Patent No. 5206352
;   APPLICANT: Roninson, Igor B.;Pastan Ira H.;Gottesman,
;Michael M.
;   TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA
;SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
;   NUMBER OF SEQUENCES: 4
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/622,836
;   FILING DATE: 24-SEP-1990
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 892,575
;   FILING DATE: 01-AUG-1986
;   APPLICATION NUMBER: 845,610
;   FILING DATE: 28-MAR-1986
;SEQ ID NO:3:
;   LENGTH: 4669
5206352-3

```
Alignment Scores:
Pred. No.:             0              Length:         4669
Score:                 3519.50        Matches:        681
Percent Similarity:    73.84%         Conservative:   245
Best Local Similarity: 54.31%         Mismatches:     265
Query Match:           57.14%         Indels:         63
DB:                    6              Gaps:           10

US-09-873-409-5 (1-1222) x 5206352-3 (1-4669)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||::::::|||   ||||||::::::::::::::||||||   ||||||||||||   |||
Db     575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy      21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
           |||   |||||||||:::|||                          :::|||:::|||
Db     635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe----------------- 43
                 ::::::      |||      ||||||
Db     695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                 |||         ||||||   |||||||||         |||:::      |||
Db     755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
           ||||||:::|||   |||   :::   |||   ||||||      :::||||||||||||
Db     815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
           |||||||||:::::::::   |||:::||||||||||||||||||         |||:::||||||
Db     875 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
           ||||||||||||:::|||   |||:::   ||||||:::      ||||||||||||||||
Db     935 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATCAAGTTATTGGTGACAAAATT 994

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
              :::   ||||||:::|||:::||||||      |||      ||||||      :::
Db     995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
           |||||||||||||||||||||   |||:::   ||||||:::      |||||||||
Db    1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
              ::::::::::::   |||   |||   |||||||||   ||||||:::||||||
Db    1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
           ||||||||||||||||||||||||:::::::||||||||||||||||||||   |||
Db    1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy     201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
           :::|||||||||:::
Db    1235 AAGAAAGAACTTGAA------------------------------------------- 1249

Qy     221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                                 ||||||   :::||||||::::::
Db    1250 ------------------------------------AGGTACAACAAAAATTTAGAAGAA 1273

Qy     241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
           ||||||   |||||||||:::   |||   :::   :::|||:::||||||   :::
Db    1274 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 1333

Qy     261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
                 :::      :::|||   ||||||||||||||||||:::|||:::|||:::
Db    1334 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 1393

Qy     281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
           ||||||      |||:::||||||   ||||||   |||||||||||||||:::
Db    1394 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1447

Qy     301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
           ::::::   :::|||   |||   |||         |||   ||||||   ||||||||||||
```

```
Db   1448 GCTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1507

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   1508 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1567

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   1568 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1627

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   1628 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1687

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   1688 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1747

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   1748 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1807

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   1808 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1867

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   1868 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1927

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   1928 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1987

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   1988 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 2047

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   2048 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 2107

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db   2108 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACC 2167

Qy    541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db   2168 ACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC 2227

Qy    561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db   2228 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATT 2287

Qy    581 TyrTyrSerLeuValMetSerGln------------------AspIleLysLysAlaAsp 594
Db   2288 TACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 2347

Qy    595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
Db   2348 GATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGT 2407

Qy    615 SerVal------------------LysSerIleLys---SerAspPheIleAspLysAla 627
Db   2408 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAG 2467

Qy    628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
Db   2468 CTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATG 2527

Qy    648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
Db   2528 AAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAAT 2587

Qy    668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
Db   2588 GGAGGCCTGCAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGA 2647

Qy    688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db   2648 ATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCC 2707

Qy    707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db   2708 CTTGGAATTATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGA 2767

Qy    727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db   2768 GAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTG 2827

Qy    747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
Db   2828 AGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGAT 2887

Qy    767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db   2888 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCA 2947

Qy    787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
Db   2948 AATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTC 3007

Qy    807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db   3008 TTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGA 3067

Qy    827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db   3068 CAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 3127

Qy    847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db   3128 GAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCT 3187

Qy    867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db   3188 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 3247

Qy    887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db   3248 TTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCC 3307

Qy    907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db   3308 TACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTT 3367

Qy    927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db   3368 GTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCC 3427

Qy    947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db   3428 AAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTAC 3487

Qy    967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db   3488 AGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTA 3547

Qy    987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db   3548 TTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAG 3607

Qy   1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
Db   3608 AAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAG 3667
```

```
Qy    1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
           ||||||:::|||   |||||||||:::   |||:::||||||   ||||||    :::
Db    3668 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATA 3727

Qy    1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
           |||   |||||||||||||||||||||:::   :::   ||||||   |||||||||:::
Db    3728 AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATC 3787

Qy    1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
           ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3788 CTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCA 3847

Qy    1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
              :::||||||      ||||||      ||||||||||||:::|||||||||   |||
Db    3848 CAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTG 3907

Qy    1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
           |||   ||||||:::|||:::||||||   ||||||   |||||||||||||||||||||
Db    3908 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAA 3967

Qy    1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
           ||||||:::|||||||||||||||||||||:::::::::|||   |||||||||||||||
Db    3968 CAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAA 4027

Qy    1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
           ||||||||||||||||||   :::||||||||||||||||||   |||||||||||||||
Db    4028 GCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCC 4087

Qy    1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
           |||   ||||||||||||:::|||:::   ||||||||||||   |||||||||||||||
Db    4088 AGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGAC 4147

Qy    1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
           ||||||||||||      ||||||::::::||||||   ||||||||||||:::||||||
Db    4148 TTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTG 4207

Qy    1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
              :::   |||||||||   :::|||:::   |||:::
Db    4208 GCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 4249
```

RESULT 7
US-08-752-447-1
; Sequence 1, Application US/08752447
; Patent No. 5994088
; GENERAL INFORMATION:
; APPLICANT: Mechetner, Eugene
; APPLICANT: Roninson, Igor B
; TITLE OF INVENTION: Methods and Reagents for Preparing and
; TITLE OF INVENTION: Using Immunoligcal Agents Specific for P-glycoprotein
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff Ltd.
; STREET: 300 South Wacker Drive, Seventh Floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/752,447
; FILING DATE: 15-NOV-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5994088nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 95,1121
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-9808
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4669 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: 1..424
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 425..4264
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: 4265..4669
US-08-752-447-1

```
Alignment Scores:
Pred. No.:              0            Length:         4669
Score:                  3503.50      Matches:        677
Percent Similarity:     73.52%       Conservative:   245
Best Local Similarity:  53.99%       Mismatches:     269
Query Match:            56.88%       Indels:         63
DB:                     2            Gaps:           10
```

US-09-873-409-5 (1-1222) x US-08-752-447-1 (1-4669)

```
Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||::::::|||   ||||||::::::::::::||||||   ||||||||||||   |||
Db     575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy      21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
           |||   |||||||||:::|||                           :::|||:::|||
Db     635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                    ::::::         |||      ||||||
Db     695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                 |||         ||||||   |||||||||         |||:::      |||
Db     755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
           ||||||:::|||   |||   :::   |||   ||||||         :::|||||||||
Db     815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
           ||||||||:::::::::      |||:::||||||||||||||||   |||:::||||||
Db     875 TTTTTTCATGCTATTATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
           ||||||||||||:::|||   |||:::   ||||||::::::||||||||||||||||||
Db     935 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAAATT 994

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
              :::   ||||||:::|||:::||||||         |||      ||||||   :::
Db     995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
           ||||||||||||||||||   |||:::   ||||||::::::      |||||||||
Db    1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
              ::::::::::::   |||   |||   |||||||||   ||||||:::||||||
Db    1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
           |||||||||||||||||||||||||::::::|||||||||||||||||||||||   |||
Db    1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy     201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
```

```
Db   1235 AAGAAAGAACTTGAA------------------------------------------ 1249

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db   1250 ----------------------------------AGGTACAACAAAAATTTAGAAGAA 1273

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db   1274 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 1333

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db   1334 CTGCTGATCTATGCATCTTATGCTGTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 1393

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db   1394 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1447

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db   1448 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1507

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   1508 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1567

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   1568 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1627

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   1628 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1687

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   1688 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1747

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   1748 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1807

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   1808 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1867

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   1868 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1927

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   1928 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1987

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   1988 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 2047

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   2048 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 2107

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db   2108 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACC 2167

Qy    541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db   2168 ACCATTGTGATAGCTCATCGTTTGCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC 2227

Qy    561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db   2228 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATT 2287

Qy    581 TyrTyrSerLeuValMetSerGln-----------------AspIleLysLysAlaAsp 594
Db   2288 TACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 2347

Qy    595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
Db   2348 GATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGT 2407

Qy    615 SerVal-----------------LysSerIleLysSerAspPheIleAspLysAlaGlu 628
Db   2408 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAACACAGAAAG 2467

Qy    629 GluSerThr---GlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
Db   2468 CTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATG 2527

Qy    648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
Db   2528 AAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAAT 2587

Qy    668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
Db   2588 GGAGGCCTGCAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGA 2647

Qy    688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db   2648 ATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCC 2707

Qy    707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db   2708 CTTGGAATTATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGA 2767

Qy    727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db   2768 GAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTG 2827

Qy    747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
Db   2828 AGTTGGTTTCATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGAT 2887

Qy    767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db   2888 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCA 2947

Qy    787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
Db   2948 AATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTC 3007

Qy    807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db   3008 TTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTTGCTGGA 3067

Qy    827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db   3068 CAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 3127

Qy    847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db   3128 GAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCT 3187

Qy    867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db   3188 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 3247

Qy    887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db   3248 TTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCC 3307

Qy    907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db   3308 TACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTT 3367
```

```
Qy     927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
              :::||||||||||||:::|||:::                |||||:::|||:::||||||
Db    3368 GTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCC 3427

Qy     947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
           |||   ||||||||:::        ::::::|||||   |||   ||||||||
Db    3428 AAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTAC 3487

Qy     967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
           |||   ||||||      |||:::|||   |||||||||:::   |||   ||||||
Db    3488 AGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTA 3547

Qy     987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
           |||   ||||||   |||||||||:::      :::|||:::||||||||||||    ::::::
Db    3548 TTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAG 3607

Qy    1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
           :::|||:::|||:::|||   ||||||||||||||||||||||||||||||   ||||||
Db    3608 AAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAG 3667

Qy    1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
           ||||||:::|||   |||||||||:::   |||:::||||||   ||||||   :::
Db    3668 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATA 3727

Qy    1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
           |||   |||||||||||||||||||||:::   :::   ||||||   |||||||||:::
Db    3728 AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATC 3787

Qy    1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
           ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3788 CTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCA 3847

Qy    1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
              :::||||||         ||||||      ||||||||||||:::|||||||||   |||
Db    3848 CAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTG 3907

Qy    1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
           |||   ||||||:::|||:::||||||   ||||||   ||||||||||||||||||||||
Db    3908 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAA 3967

Qy    1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
           ||||||:::||||||||||||   |||:::::::::|||   ||||||||||||||||||
Db    3968 CAACGCATTGCCATAGCTCGTCGCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAA 4027

Qy    1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
           ||||||||||||||||||   :::||||||||||||||||||   |||||||||||||||
Db    4028 GCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCC 4087

Qy    1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
           |||   ||||||||||||:::|||:::      ||||||||||||||||||
Db    4088 AGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGAC 4147

Qy    1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
           ||||||||||||      ||||||:::::::||||||   ||||||||||||:::||||||
Db    4148 TTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTG 4207

Qy    1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
              :::   |||||||||   :::|||:::   |||:::
Db    4208 GCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 4249
```

RESULT 8
US-09-316-167-1
; Sequence 1, Application US/09316167
; Patent No. 6365357
; GENERAL INFORMATION:
;   APPLICANT: Mechetner, Eugene
;   APPLICANT: Roninson, Igor B
;   TITLE OF INVENTION: Methods and Reagents for Preparing and
;   TITLE OF INVENTION: Using Immunoligcal Agents Specific for P-glycoprotein
;   NUMBER OF SEQUENCES: 2
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff Ltd.
;   STREET: 300 South Wacker Drive, Seventh Floor
;   CITY: Chicago
;   STATE: Illinois
;   COUNTRY: USA
;   ZIP: 60606
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/316,167
;   FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/752,447
;   FILING DATE: 15-NOV-1996
;   ATTORNEY/AGENT INFORMATION:
;   NAME: No. 6365357nan, Kevin E
;   REGISTRATION NUMBER: 35,303
;   REFERENCE/DOCKET NUMBER: 95,1121
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 312-913-0001
;   TELEFAX: 312-913-9808
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4669 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: 1..424
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 425..4264
;   FEATURE:
;   NAME/KEY: 3'UTR
;   LOCATION: 4265..4669
US-09-316-167-1

Alignment Scores:

Pred. No.:	0	Length:	4669
Score:	3503.50	Matches:	677
Percent Similarity:	73.52%	Conservative:	245
Best Local Similarity:	53.99%	Mismatches:	269
Query Match:	56.88%	Indels:	63
DB:	4	Gaps:	10

US-09-873-409-5 (1-1222) x US-09-316-167-1 (1-4669)

```
Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||:::::|||   ||||||::::::::::::||||||   ||||||||||||   |||
Db     575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy      21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
           |||   |||||||||:::|||                              :::|||:::|||
Db     635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                       ::::::      |||         ||||||
Db     695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                 |||         ||||||   |||||||||         |||:::      |||
Db     755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
           ||||||:::|||   |||   :::   |||   ||||||         :::||||||||||||
Db     815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
```

```
Db    875 TTTTTCATGCTATTATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    935 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAAATT 994

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db   1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db   1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db   1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db   1235 AAGAAAGAACTTGAA--------------------------------------------- 1249

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db   1250 -------------------------------------AGGTACAACAAAAATTTAGAAGAA 1273

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db   1274 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 1333

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db   1334 CTGCTGATCTATGCATCTTATGCTGTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 1393

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db   1394 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1447

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db   1448 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1507

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   1508 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1567

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   1568 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1627

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   1628 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1687

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   1688 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1747

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   1748 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1807

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   1808 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1867

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   1868 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1927

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   1928 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1987

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   1988 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 2047

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   2048 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 2107

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db   2108 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACC 2167

Qy    541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db   2168 ACCATTGTGATAGCTCATCGTTTGCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC 2227

Qy    561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db   2228 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATT 2287

Qy    581 TyrTyrSerLeuValMetSerGln-----------------AspIleLysLysAlaAsp 594
Db   2288 TACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 2347

Qy    595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
Db   2348 GATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGT 2407

Qy    615 SerVal------------------LysSerIleLysSerAspPheIleAspLysAlaGlu 628
Db   2408 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAACACAGAAAG 2467

Qy    629 GluSerThr---GlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
Db   2468 CTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATG 2527

Qy    648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
Db   2528 AAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAAT 2587

Qy    668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
Db   2588 GGAGGCCTGCAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGA 2647

Qy    688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db   2648 ATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCC 2707

Qy    707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db   2708 CTTGGAATTATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGA 2767

Qy    727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db   2768 GAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTG 2827

Qy    747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
Db   2828 AGTTGGTTTCATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGAT 2887

Qy    767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db   2888 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCA 2947

Qy    787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
Db   2948 AATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTC 3007
```

```
Qy     807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
           |||:::|||   |||::::::|||:::    |||::::::|||              |||
Db    3008 TTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTTGCTGGA 3067

Qy     827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
              |||   ||||||||:::||||||:::   |||||||||||||||||||||||:::
Db    3068 CAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 3127

Qy     847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
           ||||||   |||||:::|||||||||||||:::|||:::   ||||||   ||||||
Db    3128 GAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCT 3187

Qy     867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
           :::   ||||||      :::||||||:::   :::||||||   |||   |||
Db    3188 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 3247

Qy     887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
           ::::::|||:::   |||   :::||||||:::||||||            |||||||||||||||
Db    3248 TTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCC 3307

Qy     907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
           ||||||:::            |||:::   |||   :::   :::||||||:::|||:::
Db    3308 TACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTT 3367

Qy     927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
              :::||||||||||||:::|||:::               ||||||:::|||:::||||||
Db    3368 GTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCC 3427

Qy     947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
           |||      |||||||||:::         ::::::||||||   |||   |||||||||
Db    3428 AAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTAC 3487

Qy     967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
           |||   ||||||         |||:::|||   |||||||||:::      |||   ||||||
Db    3488 AGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTA 3547

Qy     987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
           |||   ||||||   |||||||||:::   :::|||:::|||||||||||||      ::::::
Db    3548 TTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAG 3607

Qy    1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
           :::|||:::|||:::|||   ||||||||||||||||||||||||||||||      ||||||
Db    3608 AAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAG 3667

Qy    1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
           ||||||:::|||   |||||||||:::   |||:::||||||   ||||||      :::
Db    3668 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATA 3727

Qy    1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
           |||   ||||||||||||||||||||:::   :::   ||||||   ||||||||||:::
Db    3728 AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATC 3787

Qy    1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
           ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3788 CTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCA 3847

Qy    1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
              :::||||||      ||||||      ||||||||||||:::|||||||||   |||
Db    3848 CAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTG 3907

Qy    1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
           |||   ||||||:::|||:::||||||   ||||||   ||||||||||||||||||||||
Db    3908 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAA 3967

Qy    1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
           ||||||:::||||||||||||      |||::::::::|||   |||||||||||||||||
Db    3968 CAACGCATTGCCATAGCTCGTCGCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAA 4027

Qy    1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
           ||||||||||||||||||      :::||||||||||||||||||   ||||||||||||||
Db    4028 GCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCC 4087

Qy    1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
           |||   ||||||||||:::|||:::   ||||||||||||   ||||||||||||||||
Db    4088 AGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGAC 4147

Qy    1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
              ||||||||||||      ||||||::::::||||||   |||||||||||:::||||||
Db    4148 TTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTG 4207

Qy    1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
              :::   |||||||||   :::|||:::   |||:::
Db    4208 GCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 4249
```

RESULT 9
US-09-120-513-1
; Sequence 1, Application US/09120513
; Patent No. 6025160
; GENERAL INFORMATION:
;   APPLICANT: Brun, Kimberly
;   APPLICANT: Chenery, Richard
;   APPLICANT: Ellens, Harma
;   APPLICANT: Field, John
;   APPLICANT: Yue, Lin
;   TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE
;   TITLE OF INVENTION: SEQUENCES ENCODING RAT MDR1B2 AND
;   TITLE OF INVENTION: SCREENING METHODS THEREOF
;   NUMBER OF SEQUENCES: 2
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: SmithKline Beecham Corporation
;   STREET: 709 Swedeland Road
;   CITY: King of Prussia
;   STATE: PA
;   COUNTRY:
;   ZIP: 19406
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/120,513
;   FILING DATE: 22-JUL-1998
;   CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: King, William T
;   REGISTRATION NUMBER: 30,954
;   REFERENCE/DOCKET NUMBER: GP50008
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 610-270-5015
;   TELEFAX: 610-270-5090
;   TELEX:
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4233 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: Genomic DNA
US-09-120-513-1

Alignment Scores:

Pred. No.:	0	Length:	4233
Score:	3494.00	Matches:	684
Percent Similarity:	72.87%	Conservative:	229
Best Local Similarity:	54.59%	Mismatches:	274
Query Match:	56.73%	Indels:	66
DB:	3	Gaps:	10

US-09-873-409-5 (1-1222) x US-09-120-513-1 (1-4233)

```
Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||   ||||||   ||||||::::::::::::|||         |||||||||:::   |||
Db     170 ATGGCTCTGGGAACTCTCGCTGCTATCATCCACGGAACCCTGCTTCCCCTCCTGATGCTG 229
```

```
Qy     21 ValLeuGlyGluMetSerAspAsnLeuIle---------------------------- 30
Db    230 GTGTTCGGATACATGACAGATAGTTTACCCAAGCAGAGACCCGCATTCTGCCGAGCGTT 289

Qy     31 ------SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu--------- 45
Db    290 ACTAATCAAAGTGAAATCAACAGTACACAGACCGTCAGCGACAGCAGTCTGGAGGAGGAC 349

Qy     46 -----------ThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    350 ATGGCCATGTACGCCTACTATTACACGGGCATTGGTGCCGGTGTGCTCATCGTTGCCTAC 409

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    410 ATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAAG 469

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    470 TTTTTCCATGCCATCATGAATCAGGAGATAGGCTGGTTTGACGTGAATGACGCTGGGGAG 529

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    530 CTCAACACCCGGCTCACAGATGACGTCTCCAAAATTAATGACGGAATTGGTGACAAACTT 589

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    590 GGAATGTTCTTTCAGTCCATAACGACATTTTCAGCCGGTTTTATAATAGGATTTATAAGT 649

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    650 GGTTGGAAGCTAACCCTTGTAATTTTGGCCGTCAGCCCTCTTATTGGGTTGTCATCTGCC 709

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    710 ATGTGGGCAAAGGTACTGACTTCATTTACTAATAAGGAACTCCAGGCTTATGCGAAAGCT 769

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    770 GGAGCAGTTGCCGAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCGTTTGGAGGACAA 829

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db    830 AAGAAGGAACTTGAA--------------------------------------------- 844

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db    845 -----------------------------------AGGTACAATAAAAATTTAGAAGAA 868

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db    869 GCTAAAAGAGTTGGCATAAAGAAAGCCATCACGGCCAACATTTCCATAGGTATTGCCTAC 928

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db    929 CTGTTGGTCTATGCGTCTTATGCACTGGCATTCTGGTATGGGACCTCCTTGGTCCTCTCA 988

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db    989 AATGAA------TATTCTATTGGACAAGTGCTTACCGTCTTCTTCTCTATTTTATTGGGG 1042

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db   1043 ACTTTCAGTATTGGACATTTAGCCCCAAACATAGAAGCCTTTGCAAATGCAAGAGGGGCA 1102

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   1103 GCCTATGAAATCTTCAAGATAATTGATAATGAGCCAAGCATTGACAGCTTCTCAACCAAG 1162

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   1163 GGACACAAACCAGACAGTATAATGGGAAATTTGGAATTTAAAAATGTTTACTTCAACTAC 1222

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   1223 CCATCACGAAGTGAAGTTAAGATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAG 1282

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   1283 ACGGTAGCCCTGGTTGGCAACAGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAG 1342

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   1343 AGGCTCTACGACCCCATAGAGGGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATC 1402

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   1403 AATGTGAGGTATCTGCGGGAAATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCC 1462

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   1463 ACCACGATTGCCGAAAACATTCGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAG 1522

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   1523 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTCATCATGAAACTGCCCCACAAATTTAAC 1582

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   1583 ACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATT 1642

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   1643 GCCCGGGCCCTGGTCCGCAACCCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTG 1702

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db   1703 GACACAGAAAGCGAAGCCGTGGTTCAGGCCGCTCTGGATAAGGCTAGAGAAGGCCGGACC 1762

Qy    541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db   1763 ACCATTGTGATAGCTCACCGCTTGTCTACAGTGCGCAATGCTGACGTCATTGCTGGTTTT 1822

Qy    561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db   1823 GATGGTGGTGTCATTGTGGAGCAAGGAAATCATGAAGAGCTCATGAAAGAGAAGGGCATT 1882

Qy    581 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet 600
Db   1883 TACTTCAAACTTGTCATGACACAG---ACTAGAGGAAATGAAATTGAACCAGGAAATAAT 1939

Qy    601 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 620
Db   1940 GCTTATGAATCCCAAAGTGACACTGGTGCCTCTGAGTTGACTTCAGAAAAATCA---AAA 1996

Qy    621 SerAspPheIleAspLysAla--------------------------------GluGlu 629
Db   1997 TCTCCTTTAATAAGGAGATCAATTCGCAGAAGTATCCACAGAAGACAAGACCAGGAGAGA 2056

Qy    630 SerThrGlnSerLysGlu---------IleSerLeuProGluValSerLeuLeuLysIle 646
Db   2057 AGACTTAGTTCGAAAGAGGATGTGGATGAAGATGTGCCTATGGTTTCCTTTTGGCAGATC 2116

Qy    647 LeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeu 666
Db   2117 CTAAAGCTAAATATTAGTGAATGGCCCTATTTAGTTGTGGGTGTACTTTGTGCTGTTATA 2176

Qy    667 AsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGly 686
Db   2177 AATGGGTGCATACAACCAGTGTTTGCCATAGTGTTTTCAAAGATTGTAGGGGTTTTTTCA 2236

Qy    687 ---AsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheVal 705
Db   2237 AGAGACGACGACCATGAAACCAAACAACGGAATTGTAACTTGTTTCCCTTCTCTTTCTG 2296

Qy    706 IleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAla 725
```

```
Db    2297 GTCATGGGAATGATTTCTTTTGTTACGTACTTCTTTCAAGGCTTCACATTTGGCAAAGCT 2356

Qy     726 GlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAsp 745
           ||||||||||||||    |||||||||::::::  ||||||:::||||||   ||||||
Db    2357 GGAGAGATCCTCACCAAGCGACTCCGATACATGGTCTTCAAATCCATGCTGCGACAGGAT 2416

Qy     746 IleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIle 765
           |||:::|||||||||:::    :::|||:::||||||   |||||||||   ||||||
Db    2417 ATAAGCTGGTTTGATGACCATAAAAACACCACTGGCTCGCTGACTACCAGGCTCGCTAGT 2476

Qy     766 AspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAla 785
           |||   :::   ::::::||||||   ||||||||||:::     |||:::||||||||||
Db    2477 GACGCTTCTAATGTTAAAGGGGCTATGGGCTCCAGGCTTGCTGTAGTTACCCAGAATGTA 2536

Qy     786 ThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeu 805
              |||:::|||        :::|||:::|||     :::|||||||||::::::|||   |||
Db    2537 GCAAACCTTGGCACAGGAATTATCTTATCCTTAGTCTATGGCTGGCAGCTTACACTTTTA 2596

Qy     806 IleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThr 825
           ::::::   |||   |||::::::   :::   |||:::||||||         ::::::
Db    2597 CTTGTAGTAATTATACCACTCATTGTCTTGGGTGGAATTATTGAAATGAAACTGTTGTCT 2656

Qy     826 GlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAla 845
           |||   |||   ||||||||||:::||||||:::   :::||||||||||||||||||||
Db    2657 GGTCAAGCCTTGAAGGACAAGAAAGAGCTAGAGATCTCTGGGAAGATCGCTACAGAAGCA 2716

Qy     846 LeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyr 865
           :::||||||   ||||||:::||||||||||||||||||:::   ||||||   ||||||
Db    2717 ATTGAAAACTTCCGCACTGTTGTCTCTTTGACTAGGGAGCAGAAGTTTGAAACTATGTAT 2776

Qy     866 GluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySer 885
           :::   ||||||   :::||||||   |||||||||   :::   |||
Db    2777 GCCCAGAGCTTGCAGATACCATACAGAAATGCTTTGAAGAAAGCACACGTCTTTGGGATC 2836

Qy     886 CysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGly 905
              :::||||||:::   |||   |||||||||:::|||||||||   ||||||||||||
Db    2837 ACCTTCGCCTTCACCCAGGCCATGATTTATTTTTCCTATGCTGCTTGTTTCCGGTTCGGT 2896

Qy     906 AlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAla 925
           |||||||||:::            ||||||   |||   :::   :::||||||:::|||
Db    2897 GCCTACTTGGTGGCACGAGAACTCATGACGTTTGAAAATGTTATGTTGGTATTTTCTGCT 2956

Qy     926 IleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLys 945
           :::   :::||||||||||||   |||   |||        ||||||:::|||:::|||
Db    2957 GTTGTCTTTGGTGCCATGGCAGCAGGGAATACCAGTTCATTCGCTCCTGACTACGCGAAG 3016

Qy     946 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSer 965
           ||||||   |||:::|||:::      ::::::||||||   |||   |||||||||
Db    3017 GCCAAAGTCTCAGCATCCCACATCATCAGGATCATTGAGAAAATCCCCGAGATTGACAGC 3076

Qy     966 ArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 985
              |||   ||||||   ||||||:::      |||||||||::::::|||      |||
Db    3077 TACAGCACGGAGGGCTTGAAGCCTAATTGGTTAGAAGGAAATGTGAAATTTAATGGAGTC 3136

Qy     986 SerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIle 1005
              |||   ||||||   ||||||::::::   :::|||:::||||||||      :::
Db    3137 ATGTTCAACTATCCCACCCGACCCAACATCCCAGTGCTTCAGGGACTGAGCTTCGAGGTG 3196

Qy    1006 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 1025
           ::::::|||:::|||:::      |||||||||||||||||||||||||||   |||
Db    3197 AAGAAGGGGCAAACGCTTCGCCTGGTGGGCAGCAGTGGCTGCGGGAAGAGTACAGTGGTC 3256

Qy    1026 GlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAsp 1045
           |||||||||:::|||   |||:::|||   |||   |||      ||||||   :::
Db    3257 CAGCTGCTCGAGCGCTTCTACAACCCCATGGCTGGAACAGTGTTTCTAGATGGCAAAGAA 3316

Qy    1046 AlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 1065
              |||:::||||||||||||||||||||:::   :::   ||||||   |||||||||
Db    3317 ATAAAACAACTCAACGTCCAGTGGCTCCGCGCCCACCTGGGCATTGTGTCCCAGGAGCCC 3376

Qy    1066 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValVal 1085
           :::||||||:::|||||||||   ||||||||||||||||||||||||||||||||||||
Db    3377 ATCCTGTTTGACTGCAGCATCACCGAGAACATCGCCTACGGAGACAACAGCCGTGTCGTG 3436

Qy    1086 ProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGly 1105
              :::||||||   ||||||   ||||||||||||   ||||||:::
Db    3437 TCTCATGAGGAGATCGTGAGGGCCGCCAGGGAGGCCAACATCCACCAGTTCATCGACTCA 3496

Qy    1106 LeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGln 1125
           |||||||||||||||||||||:::||||||   ||||||   ||||||||||||||||||
Db    3497 CTGCCTGAGAAATACAACACCAGAGTGGGAGACAAAGGGACTCAGCTGTCGGGCGGGCAG 3556

Qy    1126 LysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAsp 1145
           |||||||||:::||||||||||||||||||:::::::::|||   |||||||||||||||
Db    3557 AAGCAGCGCATCGCCATCGCGCGCGCCCTCGTCAGACAGCCTCACATCTTACTTCTGGAT 3616

Qy    1146 GluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLys 1165
           |||||||||||||||||||   :::|||||||||||||||||   ||||||||||||
Db    3617 GAAGCGACATCAGCTCTGGATACGGAGAGTGAAAAGGTCGTCCAGGAAGCGCTGGACAAA 3676

Qy    1166 AlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla 1185
           ||||||   ||||||||||||:::|||:::      ||||||   ||||||||||||
Db    3677 GCCAGGGAAGGCCGCACCTGCATTGTGATCGCGCACCGCCTGTCCACCATCCAGAACGCA 3736

Qy    1186 AspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1205
           |||||||||||||||:::   ||||||::::::||||||   ||||||||||||:::|||
Db    3737 GACTTGATCGTGGTGATTCAGAACGGCCAGGTCAAGGAGCACGGCACCCACCAGCAGCTG 3796

Qy    1206 LeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 1218
           |||   :::   |||||||||   :::|||   |||
Db    3797 CTGGCCCAGAAAGGCATCTATTTCTCGATGGTTCAGGCT 3835

RESULT 10
US-09-450-105-1
; Sequence 1, Application US/09450105
; Patent No. 6169166
; GENERAL INFORMATION:
;  APPLICANT: Kimberly Anne Brun
;  APPLICANT: Richard James Chenery
;  APPLICANT: Harma Ellens
;  APPLICANT: John Anthony Feild
;  APPLICANT: Lin Yue
;  TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
;  TITLE OF INVENTION: ENCODING RAT MDR1B2 AND SCREENING METHODS THEREOF
;  FILE REFERENCE: GP-50008-D1
;  CURRENT APPLICATION NUMBER: US/09/450,105
;  CURRENT FILING DATE: 1999-11-29
;  EARLIER APPLICATION NUMBER: 09/120,513
;  EARLIER FILING DATE: 1998-07-22
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 4233
;   TYPE: DNA
;   ORGANISM: HOMO SAPIENS
US-09-450-105-1

Alignment Scores:
Pred. No.:              0              Length:        4233
Score:                  3494.00        Matches:       684
Percent Similarity:     72.87%         Conservative:  229
Best Local Similarity:  54.59%         Mismatches:    274
Query Match:            56.73%         Indels:        66
DB:                     4              Gaps:          10

US-09-873-409-5 (1-1222) x US-09-450-105-1 (1-4233)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||   ||||||   ||||||::::::::::::|||      |||||||||:::   |||
Db     170 ATGGCTCTGGGAACTCTCGCTGCTATCATCCACGGAACCCTGCTTCCCCTCCTGATGCTG 229

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
           |||   |||   |||:::|||:::
Db     230 GTGTTCGGATACATGACAGATAGTTTTACCCAAGCAGAGACCCGCATTCTGCCGAGCGTT 289

Qy      31 ------SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu--------- 45
```

```
Db     290 ACTAATCAAAGTGAAATCAACAGTACACAGACCGTCAGCGACAGCAGTCTGGAGGAGGAC 349

Qy      46 ------------ThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db     350 ATGGCCATGTACGCCTACTATTACACGGGCATTGGTGCCGGTGTGCTCATCGTTGCCTAC 409

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db     410 ATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAAG 469

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db     470 TTTTTCCATGCCATCATGAATCAGGAGATAGGCTGGTTTGACGTGAATGACGCTGGGGAG 529

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db     530 CTCAACACCCGGCTCACAGATGACGTCTCCAAAATTAATGACGGAATTGGTGACAAACTT 589

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db     590 GGAATGTTCTTTCAGTCCATAACGACATTTTCAGCCGGTTTTATAATAGGATTTATAAGT 649

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db     650 GGTTGGAAGCTAACCCTTGTAATTTTGGCCGTCAGCCCTCTTATTGGGTTGTCATCTGCC 709

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db     710 ATGTGGGCAAAGGTACTGACTTCATTTACTAATAAGGAACTCCAGGCTTATGCGAAAGCT 769

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db     770 GGAGCAGTTGCCGAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCGTTTGGAGGACAA 829

Qy     201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db     830 AAGAAGGAACTTGAA--------------------------------------------- 844

Qy     221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db     845 ------------------------------------AGGTACAATAAAAATTTAGAAGAA 868

Qy     241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db     869 GCTAAAAGAGTTGGCATAAAGAAAGCCATCACGGCCAACATTTCCATAGGTATTGCCTAC 928

Qy     261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db     929 CTGTTGGTCTATGCGTCTTATGCACTGGCATTCTGGTATGGGACCTCCTTGGTCCTCTCA 988

Qy     281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db     989 AATGAA------TATTCTATTGGACAAGTGCTTACCGTCTTCTTCTCTATTTTATTGGGG 1042

Qy     301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db    1043 ACTTTCAGTATTGGACATTTAGCCCCAAACATAGAAGCCTTTGCAAATGCAAGAGGGGCA 1102

Qy     321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db    1103 GCCTATGAAATCTTCAAGATAATTGATAATGAGCCAAGCATTGACAGCTTCTCAACCAAG 1162

Qy     341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db    1163 GGACACAAACCAGACAGTATAATGGGAAATTTGGAATTTAAAAATGTTTACTTCAACTAC 1222

Qy     361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db    1223 CCATCACGAAGTGAAGTTAAGATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAG 1282

Qy     381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db    1283 ACGGTAGCCCTGGTTGGCAACAGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAG 1342

Qy     401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db    1343 AGGCTCTACGACCCCATAGAGGGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATC 1402

Qy     421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db    1403 AATGTGAGGTATCTGCGGGAAATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCC 1462

Qy     441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db    1463 ACCACGATTGCCGAAAACATTCGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAG 1522

Qy     461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db    1523 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTCATCATGAAACTGCCCCACAAATTTAAC 1582

Qy     481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db    1583 ACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATT 1642

Qy     501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db    1643 GCCCGGGCCCTGGTCCGCAACCCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTG 1702

Qy     521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db    1703 GACACAGAAAGCGAAGCCGTGGTTCAGGCCGCTCTGGATAAGGCTAGAGAAGGCCGGACC 1762

Qy     541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db    1763 ACCATTGTGATAGCTCACCGCTTGTCTACAGTGCGCAATGCTGACGTCATTGCTGGTTTT 1822

Qy     561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db    1823 GATGGTGGTGTCATTGTGGAGCAAGGAAATCATGAAGAGCTCATGAAAGAGAAGGGCATT 1882

Qy     581 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet 600
Db    1883 TACTTCAAACTTGTCATGACACAG---ACTAGAGGAAATGAAATTGAACCAGGAAATAAT 1939

Qy     601 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 620
Db    1940 GCTTATGAATCCCAAAGTGACACTGGTGCCTCTGAGTTGACTTCAGAAAAATCA---AAA 1996

Qy     621 SerAspPheIleAspLysAla---------------------------------GluGlu 629
Db    1997 TCTCCTTTAATAAGGAGATCAATTCGCAGAAGTATCCACAGAAGACAAGACCAGGAGAGA 2056

Qy     630 SerThrGlnSerLysGlu---------IleSerLeuProGluValSerLeuLeuLysIle 646
Db    2057 AGACTTAGTTCGAAAGAGGATGTGGATGAAGATGTGCCTATGGTTTCCTTTTGGCAGATC 2116

Qy     647 LeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeu 666
Db    2117 CTAAAGCTAAATATTAGTGAATGGCCCTATTTAGTTGTGGGTGTACTTTGTGCTGTTATA 2176

Qy     667 AsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGly 686
Db    2177 AATGGGTGCATACAACCAGTGTTTGCCATAGTGTTTTCAAAGATTGTAGGGGTTTTTTCA 2236

Qy     687 ---AsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheVal 705
Db    2237 AGAGACGACGACCATGAAACCAAACAACGGAATTGTAACTTGTTTTCCCTTCTCTTTCTG 2296

Qy     706 IleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAla 725
Db    2297 GTCATGGGAATGATTTCTTTTGTTACGTACTTCTTTCAAGGCTTCACATTTGGCAAAGCT 2356

Qy     726 GlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAsp 745
Db    2357 GGAGAGATCCTCACCAAGCGACTCCGATACATGGTCTTCAAATCCATGCTGCGACAGGAT 2416
```

```
Qy     746 IleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIle 765
Db    2417 ATAAGCTGGTTTGATGACCATAAAAACACCACTGGCTCGCTGACTACCAGGCTCGCTAGT 2476

Qy     766 AspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAla 785
Db    2477 GACGCTTCTAATGTTAAAGGGGCTATGGCTCCAGGCTTGCTGTAGTTACCCAGAATGTA 2536

Qy     786 ThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeu 805
Db    2537 GCAAACCTTGGCACAGGAATTATCTTATCCTTAGTCTATGGCTGGCAGCTTACACTTTTA 2596

Qy     806 IleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThr 825
Db    2597 CTTGTAGTAATTATACCACTCATTGTCTTGGGTGGAATTATTGAAATGAAACTGTTGTCT 2656

Qy     826 GlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAla 845
Db    2657 GGTCAAGCCTTGAAGGACAAGAAAGAGCTAGAGATCTCTGGGAAGATCGCTACAGAAGCA 2716

Qy     846 LeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyr 865
Db    2717 ATTGAAAACTTCCGCACTGTTGTCTCTTTGACTAGGGAGCAGAAGTTTGAAACTATGTAT 2776

Qy     866 GluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySer 885
Db    2777 GCCCAGAGCTTGCAGATACCATACAGAAATGCTTTGAAGAAAGCACACGTCTTTGGGATC 2836

Qy     886 CysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGly 905
Db    2837 ACCTTCGCCTTCACCCAGGCCATGATTTATTTTTCCTATGCTGCTTGTTTCCGGTTCGGT 2896

Qy     906 AlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAla 925
Db    2897 GCCTACTTGGTGGCACGAGAACTCATGACGTTTGAAAATGTTATGTTGGTATTTTCTGCT 2956

Qy     926 IleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLys 945
Db    2957 GTTGTCTTTGGTGCCATGGCAGCAGGGAATACCAGTTCATTCGCTCCTGACTACGCGAAG 3016

Qy     946 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSer 965
Db    3017 GCCAAAGTCTCAGCATCCCACATCATCAGGATCATTGAGAAAATCCCCGAGATTGACAGC 3076

Qy     966 ArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 985
Db    3077 TACAGCACGGAGGGCTTGAAGCCTAATTGGTTAGAAGGAAATGTGAAATTTAATGGAGTC 3136

Qy     986 SerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIle 1005
Db    3137 ATGTTCAACTATCCCACCCGACCCAACATCCCAGTGCTTCAGGGACTGAGCTTCGAGGTG 3196

Qy    1006 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 1025
Db    3197 AAGAAGGGGCAAACGCTTCGCCTGGTGGGCAGCAGTGGCTGCGGGAAGAGTACAGTGGTC 3256

Qy    1026 GlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAsp 1045
Db    3257 CAGCTGCTCGAGCGCTTCTACAACCCCATGGCTGGAACAGTGTTTCTAGATGGCAAAGAA 3316

Qy    1046 AlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 1065
Db    3317 ATAAAACAACTCAACGTCCAGTGGCTCCGCGCCCACCTGGGCATTGTGTCCCAGGAGCCC 3376

Qy    1066 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValVal 1085
Db    3377 ATCCTGTTTGACTGCAGCATCACCGAGAACATCGCCTACGGAGACAACAGCCGTGTCGTG 3436

Qy    1086 ProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGly 1105
Db    3437 TCTCATGAGGAGATCGTGAGGGCCGCCAGGGAGGCCAACATCCACCAGTTCATCGACTCA 3496

Qy    1106 LeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGln 1125
Db    3497 CTGCCTGAGAAATACAACACCAGAGTGGGAGACAAAGGGACTCAGCTGTCGGGCGGGCAG 3556

Qy    1126 LysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAsp 1145
Db    3557 AAGCAGCGCATCGCCATCGCGCGCGCCCTCGTCAGACAGCCTCACATCTTACTTCTGGAT 3616

Qy    1146 GluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLys 1165
Db    3617 GAAGCGACATCAGCTCTGGATACGGAGAGTGAAAAGGTCGTCCAGGAAGCGCTGGACAAA 3676

Qy    1166 AlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla 1185
Db    3677 GCCAGGGAAGGCCGCACCTGCATTGTGATCGCGCACCGCCTGTCCACCATCCAGAACGCA 3736

Qy    1186 AspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1205
Db    3737 GACTTGATCGTGGTGATTCAGAACGGCCAGGTCAAGGAGCACGGCACCCACCAGCAGCTG 3796

Qy    1206 LeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 1218
Db    3797 CTGGCCCAGAAAGGCATCTATTTCTCGATGGTTCAGGCT 3835
```

```
RESULT 11
US-08-583-276-18
; Sequence 18, Application US/08583276
; Patent No. 5837536
;  GENERAL INFORMATION:
;    APPLICANT:  McDonagh, Kevin T.
;    APPLICANT:  Nienhuis, Arthur
;    APPLICANT:  Tolstoshev, Paul
;    TITLE OF INVENTION:  IMPROVED EXPRESSION OF HUMAN
;    TITLE OF INVENTION:  MULTIDRUG RESISTANCE GENES AND IMPROVED
;    TITLE OF INVENTION:  SELECTION OF CELLS TRANSDUCED WITH SUCH GENES
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carella, Byrne, Bain, Gilfillan,
;      ADDRESSEE:  Cecchi & Stewart
;      STREET:  6 Becker Farm Road
;      CITY:  Roseland
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch diskette
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  DW4.V2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/583,276
;      FILING DATE:  05-JAN-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/332,444
;      FILING DATE:  31-OCT-1994
;      APPLICATION NUMBER:  07/887,712
;      FILING DATE:  22-MAY-1992
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4669 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  singular
;      TOPOLOGY:  linear
;    MOLECULE TYPE:
;      DESCRIPTION:  Genomic DNA
US-08-583-276-18

Alignment Scores:
Pred. No.:               0            Length:         4669
Score:                   3473.50      Matches:        673
Percent Similarity:      73.34%       Conservative:   246
Best Local Similarity:   53.71%       Mismatches:     271
```

```
Query Match:              56.40%      Indels:           63
DB:                       2           Gaps:             10

US-09-873-409-5 (1-1222) x US-08-583-276-18 (1-4669)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            |||:::::|||   |||||:::::::::::|||||   |||||||||||   |||
Db      575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy       21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
            |||  ||||||||:::|||                             :::|||:::|||
Db      635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy       32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                     ::::::        |||        ||||||
Db      695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy       44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                  |||          ||||||  |||||||||        |||:::      |||
Db      755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy       62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
            ||||||:::|||   |||   :::   |||   ||||||      :::||||||||||||
Db      815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy       82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
            |||||||||:::::::::      |||:::||||||||||||||||      |||:::|||:::
Db      875 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGCAG 934

Qy      102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
            ||||||||||||:::|||   |||:::      ||||||:::::::   ||||||||||||||
Db      935 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAACTTATTGGTGACAAAATT 994

Qy      121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
               :::   ||||||:::|||:::||||||        |||      ||||||      :::
Db      995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy      141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
            ||||||||||||||||||||||   |||:::   ||||||::::::      |||||||||
Db     1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy      161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
               ::::::::::::   |||   |||   |||||||||   ||||||:::||||||
Db     1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy      181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
            ||||||||||||||||||||||||:::::::||||||||||||||||||||      |||
Db     1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy      201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
            :::|||||||||||:::
Db     1235 AAGAAAGAACTTGAA--------------------------------------------- 1249

Qy      221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                                ||||||   :::||||||::::::
Db     1250 -------------------------------------AGGTACAACAAAAATTTAGAAGAA 1273

Qy      241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
            ||||||      |||||||||:::   |||   :::   :::|||:::||||||      :::
Db     1274 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 1333

Qy      261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
               :::      :::|||   ||||||||||||||||||||:::|||:::|||:::
Db     1334 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 1393

Qy      281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
            ||||||      |||:::||||||   ||||||   ||||||||||||||||:::
Db     1394 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1447

Qy      301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
            ::::::   :::|||   |||   |||      |||   ||||||   ||||||||||||
Db     1448 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1507

Qy      321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
            |||:::      ||||||::::::|||||   ||||||||||||||:::::::|||      :::
Db     1508 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1567

Qy      341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
            |||:::|||||||:::::::|||:::|||   :::||||||:::||||||   |||:::|||
Db     1568 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1627

Qy      361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
            ||||||||      :::||||||||||||||||||||||||:::::::::::||||||:::
Db     1628 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1687

Qy      381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
            ||||||   |||||||||   :::|||   ||||||||||||   |||||||||:::|||
Db     1688 ACGGTCCCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1747

Qy      401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
            ||||||||||||||||   :::|||   :::   |||         |||||||||   :::
Db     1748 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTCATGGACAGGATATTAGGACCATA 1807

Qy      421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
            ||||||||      |||:::   ||||||||||||||||||||||||||||||
Db     1808 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1867

Qy      441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
            |||||||||:::   ||||||:::||||||||||::::::||||||   :::|||:::|||
Db     1868 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1927

Qy      461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
            :::|||   :::||||||||||||||||||||||||||||:::   |||:::||||||:::
Db     1928 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1987

Qy      481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
            |||||||||||||||:::|||   |||:::||||||||||||||||||||||||||||
Db     1988 ACCCTGGTTGGAGAGAGAGGGCCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 2047

Qy      501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
            ||||||||||||||||||||||||||||:::||||||||||||||||||||||||||||
Db     2048 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 2107

Qy      521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
            |||:::||||||::::::   ||||||   ||||||:::||||||   ||||||||||||
Db     2108 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACC 2167

Qy      541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
            |||||||||:::||||||||||||||||||:::|||:::|||||||:::|||
Db     2168 ACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC 2227

Qy      561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
               ||||||::::::   |||||||||   |||   |||||||||   ::::::|||:::
Db     2228 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATT 2287

Qy      581 TyrTyrSerLeuValMetSerGln------------------AspIleLysLysAlaAsp 594
            |||:::   ||||||   |||                     :::::::::   :::
Db     2288 TACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 2347

Qy      595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
            ::::::   :::|||         :::      |||      :::|||
Db     2348 GATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCACGT 2407

Qy      615 SerVal------------------LysSerIleLys---SerAspPheIleAspLysAla 627
               :::                  :::|||::::::   |||         |||:::
Db     2408 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAACCCCAAGACAGAAAG 2467

Qy      628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
               :::   ::::::   :::   |||:::|||   ||||||      :::|||:::
Db     2468 CTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATG 2527

Qy      648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
            |||||||||      |||||||||:::   |||:::|||      :::::::::|||
Db     2528 AAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAAT 2587

Qy      668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
```

```
Db   2588 GGAGGCCTGCAACCAGCATTTGCAATAATATTTCAAAGATTATAGGGGTTTTTACAAGA 2647

Qy    688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db   2648 ATTGATGATCCTGAAACAAAACGACAGAATACTAACTTGTTTTCACTATTGTTTCTAGCC 2707

Qy    707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db   2708 CTTGGAATTATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGA 2767

Qy    727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db   2768 GAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTG 2827

Qy    747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
Db   2828 AGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGAT 2887

Qy    767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db   2888 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCA 2947

Qy    787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
Db   2948 ATTCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTC 3007

Qy    807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db   3008 TTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGA 3067

Qy    827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db   3068 CAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 3127

Qy    847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db   3128 GAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCT 3187

Qy    867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db   3188 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 3247

Qy    887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db   3248 TTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCC 3307

Qy    907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db   3308 TACTTGGTGGCACATAAACTAATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTT 3367

Qy    927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db   3368 GTCTTTGGTGCCATGGCCGTGGGGCAAGTGAGTTCATTTGCTCCTGACTATGCCAAAGCC 3427

Qy    947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db   3428 AAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTAC 3487

Qy    967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db   3488 AGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTA 3547

Qy    987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db   3548 TTCAACTATCCCACCCGACCGGACATCCCAGTCCTTCAGGGACTGAGCCTGGAGGTGAAG 3607

Qy   1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
Db   3608 AAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAG 3667

Qy   1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
Db   3668 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATA 3727

Qy   1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
Db   3728 AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATC 3787

Qy   1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db   3788 CTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGACACAACAGCCGGGTGGTGTCA 3847

Qy   1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db   3848 CAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTG 3907

Qy   1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
Db   3908 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAA 3967

Qy   1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
Db   3968 CAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAA 4027

Qy   1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
Db   4028 GCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAACCC 4087

Qy   1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db   4088 AGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGAC 4147

Qy   1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
Db   4148 TTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGCAGCATGGCACGCATCAGCAGCTGCTG 4207

Qy   1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGln 1219
Db   4208 GCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAG 4246
```

```
RESULT 12
US-08-461-823-1
; Sequence 1, Application US/08461823
; Patent No. 5593840
;  GENERAL INFORMATION:
;    APPLICANT:  Bhatnagar, Satish K.
;    APPLICANT:  George Jr., Albert L.
;    APPLICANT:  Nazarenko, Irina
;    TITLE OF INVENTION:  AMPLIFICATION OF NUCLEIC ACID SEQUENCES
;    NUMBER OF SEQUENCES:  27
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  OncorPharm, Inc.
;      STREET:  200 Perry Parkway
;      CITY:  Gaithersburg
;      STATE:  Maryland
;      COUNTRY:  USA
;      ZIP:  20877
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/461,823
;      FILING DATE:  05-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/168,621
;      FILING DATE:  16-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/010,433
;      FILING DATE:  27-JAN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Karta, Glenn E.
;      REGISTRATION NUMBER:  30,649
```

```
;      REFERENCE/DOCKET NUMBER:  PA-0012 CIP 2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  301 527-2058
;       TELEFAX:  301 208-6997
;   INFORMATION FOR SEQ ID NO:  1:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  2726 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  double
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
;     HYPOTHETICAL:  NO
;     ANTI-SENSE:  NO
US-08-461-823-1

Alignment Scores:
Pred. No.:                1.31e-245     Length:          2726
Score:                    2280.00       Matches:         435
Percent Similarity:       76.55%        Conservative:    159
Best Local Similarity:    56.06%        Mismatches:      168
Query Match:              37.02%        Indels:          14
DB:                       1             Gaps:            4

US-09-873-409-5 (1-1222) x US-08-461-823-1 (1-2726)

Qy     459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
           :::|||:::|||   :::||||||||||||||||||||||||||:::    |||:::|||
Db       2 ATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAA 61

Qy     479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
           |||:::||||||||||||||||:::|||||||||:::||||||||||||||||||||||||
Db      62 TTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATC 121

Qy     499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
           ||||||||||||||||||||||||||||||||||||:::|||||||||||||||||||||
Db     122 GCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCA 181

Qy     519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 538
           |||||||||:::||||||::::::   ||||||   ||||||:::||||||   ||||||
Db     182 GCCTTGGACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGT 241

Qy     539 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558
           |||||||||||||||:::||||||||||||||||||||||:::|||:::||||||:::|||
Db     242 CGGACCACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCT 301

Qy     559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578
                    ||||||::::::   ||||||||   |||   |||||||||    ::::::
Db     302 GGTTTCGATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAA 361

Qy     579 GlyLeuTyrTyrSerLeuValMetSerGln------------------AspIleLysLys 592
           |||:::|||:::   ||||||   |||                       :::::::::
Db     362 GGCATTTACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAAT 421

Qy     593 AlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPro 612
           |||   ::::::   :::|||            :::   |||      :::|||
Db     422 GCAGCTGATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGA 481

Qy     613 LeuHisSerVal------------------LysSerIleLys---SerAspPheIleAsp 625
                    :::                  :::|||::::::   |||         |||
Db     482 TCCAGTCTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGAC 541

Qy     626 LysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLys 645
           :::         :::   ::::::   :::   |||:::|||   ||||||         :::
Db     542 AGAAAGCTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGG 601

Qy     646 IleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerVal 665
           |||:::|||||||||      |||||||||:::   |||:::|||         ::::::
Db     602 ATTATGAAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATT 661

Qy     666 LeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPhe 685
           :::||||||   :::      |||      |||:::||||||||||:::||||||   :::|||
Db     662 ATAAATGGAGGCCTGCAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTT 721

Qy     686 GlyAsn---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 704
                    :::|||      |||   :::      ::::::   ::::::|||::::::|||
Db     722 ACAAGAATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTT 781

Qy     705 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 724
           :::   ||||||:::|||   |||::::::::::|||:::||||||         :::|||:::
Db     782 CTAGCCCTTGGAATTATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAA 841

Qy     725 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 744
           ||||||||||||||||||   |||||||||::::::   |||::::::||||||   |||
Db     842 GCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAG 901

Qy     745 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 764
           |||::::::|||||||||:::   :::|||:::||||||   |||||||||   ||||||
Db     902 GATGTGAGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCC 961

Qy     765 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 784
              |||   ||||||::::::||||||   |||||||||:::   |||:::|||||||||
Db     962 AATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAAT 1021

Qy     785 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 804
                    |||:::|||      :::||||||||||||||||||||||||:::::::|||
Db    1022 ATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTG 1081

Qy     805 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 824
           |||:::|||:::|||   |||::::::|||:::   |||::::::|||         :::
Db    1082 TTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTG 1141

Qy     825 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 844
           :::|||   |||   |||||||||:::||||||:::   ||||||||||||||||||||
Db    1142 TCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAA 1201

Qy     845 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 864
           |||:::||||||   ||||||:::||||||||||||:::|||:::   ||||||   |||
Db    1202 GCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATG 1261

Qy     865 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 884
           |||   :::   ||||||   :::   :::|||||||||   |||   |||   |||
Db    1262 TATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGA 1321

Qy     885 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 904
              ::::::|||:::   |||   :::||||||:::||||||      |||||||||
Db    1322 ATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTT 1381

Qy     905 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 924
           ||||||||||||:::               |||:::   |||   :::   :::||||||:::
Db    1382 GGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCA 1441

Qy     925 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 944
           |||:::   :::||||||||||||:::|||:::               ||||||:::|||:::
Db    1442 GCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCC 1501

Qy     945 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 964
           |||||||||      |||||||||:::      ::::::||||||   |||   ||||||
Db    1502 AAAGCCAAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGAC 1561

Qy     965 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 984
           |||   |||   ||||||      ||||||:::|||   |||||||||:::   |||   |||
Db    1562 AGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAA 1621

Qy     985 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 1004
           |||   |||   ||||||   |||||||||:::   :::|||:::||||||||||||
Db    1622 GTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAG 1681

Qy    1005 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 1024
           ::::::::|||:::|||:::|||   |||||||||||||||||||||||||||
Db    1682 GTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTG 1741

Qy    1025 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 1044
           ||||||||||||:::|||   |||||||||:::   |||:::||||||   ||||||
Db    1742 GTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAA 1801

Qy    1045 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 1064
```

```
                 :::   |||   ||||||||||||||||||||:::   :::   ||||||   ||||||
Db    1802 GAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAG 1861

Qy    1065 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 1084
           |||:::||||||:::||||||||||||||||||||||||||||||||||||||||||||||
Db    1862 CCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTG 1921

Qy    1085 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 1104
           |||      :::||||||      ||||||      ||||||||||||:::||||||||
Db    1922 GTGTCACAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAG 1981

Qy    1105 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 1124
              ||||||   ||||||:::|||:::||||||   ||||||   ||||||||||||||
Db    1982 TCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGC 2041

Qy    1125 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 1144
           ||||||||||||:::||||||||||||||||||:::::::::|||   ||||||||||||
Db    2042 CAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTG 2101

Qy    1145 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1164
           ||||||||||||||||||||||||      :::||||||||||||||||||   ||||||||
Db    2102 GATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGAC 2161

Qy    1165 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1184
           |||||||||   ||||||||||||:::|||:::   ||||||||||||   |||||||||
Db    2162 AAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAAT 2221

Qy    1185 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1204
           ||||||||||||||||||         ||||||::::::||||||   ||||||||||||:::
Db    2222 GCAGACTTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAG 2281

Qy    1205 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
           ||||||      :::   |||||||||   :::|||:::   |||:::
Db    2282 CTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 2329
```

RESULT 13
US-08-996-545-1
; Sequence 1, Application US/08996545
; Patent No. 5928898
; GENERAL INFORMATION:
; APPLICANT: Skatrud, Paul L.
; APPLICANT: de Waard, Maarten A.
; APPLICANT: Peery, Robert B.
; APPLICANT: Andrade, Alan C.
; TITLE OF INVENTION: Multiple Drug Resistance Gene atrD of
; TITLE OF INVENTION: Aspergillus Nidulans
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/996,545
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Webster, Thomas D.
; REGISTRATION NUMBER: 39,872
; REFERENCE/DOCKET NUMBER: X-11766
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3334
; TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4002 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4002
US-08-996-545-1

```
Alignment Scores:
Pred. No.:               6.01e-243     Length:          4002
Score:                   2259.00       Matches:         494
Percent Similarity:      58.41%        Conservative:    232
Best Local Similarity:   39.74%        Mismatches:      446
Query Match:             36.68%        Indels:          72
DB:                      2             Gaps:            12

US-09-873-409-5 (1-1222) x US-08-996-545-1 (1-4002)

Qy      15 LeuProLeuMetProLeuValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu 34
           :::||||||:::|||      |||                              |||
Db     368 GTGCCATTGCTGCCGCGTCGACTTTCCAGA------------------GGATAATGTTAT 409

Qy      35 ValGlnThrAsnThrTyrSerPhePhe------------ArgLeuThrLeuTyrTyrVal 50
           :::::::::   |||      |||:::            :::      ||||||:::|||
Db     410 ATCAAATCTCGTACGAC-GAGTTCTATGATGAATTGACCAAGAACGTACTGTACTTCGTA 468

Qy      51 GlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThr 70
              :::|||:::         :::      |||:::            :::|||   |||
Db     469 TACCTCGGTATCGGCGAGTTTGTCACTGTCTATGTTAGTACTGTTGGCTTCATCTATACC 528

Qy      71 AlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAsp 90
                          |||::::::||||||:::   :::      |||:::|||   |||:::
Db     529 GGAGAACACGCCACGCAGAAGATCCGCGAGTATTACCTTGAGTCTATCCTGCGCCAGAAC 588

Qy      91 IleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMetThr---AspIle 109
           ||||||:::||||||            ||||||:::   ||||||:::|||   |||
Db     589 ATTGGCTATTTTGATAAACTCGGTGCCGGGGAAGTGACCACCCGTATAACAGCCGATACA 648

Qy     110 AspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThr 129
           :::   |||   |||||||||   :::|||:::   |||            ::::::|||
Db     649 AACCTTATCCAGGATGGCATTTCGGAGAAGGTCGGTCTCACTTTGACTGCCCTGGCGACA 708

Qy     130 PheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeu 149
           |||                  :::      ||||||   |||||||||   |||:::
Db     709 TTCGTGACAGCATTCATTATCGCCTACGTCAAATACTGGAAGTTGGCTCTAATTTGCAGC 768

Qy     150 SerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeu 169
           ||||||         ::::::   :::               |||:::   :::|||
Db     769 TCAACAATTGTGGCCCTCGTTCTCACCATGGGCGGTGGTTCTCAGTTTATCATCAAGTAC 828

Qy     170 ThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSer 189
           :::   |||   |||   :::|||         |||   ||||||||||||||:::|||
Db     829 AGCAAAAAGTCGCTTGACAGCTACGGTGCAGGCGGCACTGTTGCGGAAGAGGTCATCAGC 888

Qy     190 SerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeu 209
           |||||||||         ||||||      |||:::|||         :::
Db     889 TCCATCAGAAATGCCACAGCGTTTGGCACCCAAGACAAGCTTGCGAAG------------ 936

Qy     210 LeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu 229

Db     936 ------------------------------------------------------------ 936

Qy     230 ***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThr 249
                       :::|||      :::|||   :::|||:::   :::|||   |||
Db     937 ---------CAGTATGAGGTCCACTTAGACGAAGCTGAGAAATGGGGAACAAAGAACCAG 987

Qy     250 IleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeu 269
           |||               :::||||||::::::      |||         |||||||||
Db     988 ATTGTCATGGGTTTCATGATTGGCGCCATGTTTGGCCTTATGTACTCGAACTACGGTCTT 1047
```

```
Qy    270 AlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThr 289
Db   1048 GGCTTCTGGATGGGTTCTCGTTTCCTGGTAGATGGT------GCAGTCGATGTGGGTGAT 1101

Qy    290 ValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValPro 309
Db   1102 ATTCTCACAGTTCTCATGGCCATCTTGATCGGATCGTTCTCCTTGGGGAACGTTAGTCCA 1161

Qy    310 HisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAsp 329
Db   1162 AATGCTCAAGCATTTACAAACGCTGTGGCCGCGGCCGCAAAGATATTTGGAACGATCGAT 1221

Qy    330 LysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGly 349
Db   1222 CGCCAGTCCCCATTAGATCCATATTCGAACGAAGGGAAGACGCTCGACCATTTTGAGGGC 1281

Qy    350 ThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeu 369
Db   1282 CACATTGAGTTACGCAATGTCAAGCATATTTACCCATCTAGACCCGAGGTCACCGTCATG 1341

Qy    370 LysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGly 389
Db   1342 GAGGATGTTTCTCTGTCAATGCCCGCTGGAAAAACAACCGCTTTAGTCGGCCCCTCTGGC 1401

Qy    390 SerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPhe 409
Db   1402 TCTGGAAAAAGTACGGTGGTCGGCTTGGTTGAGCGATTCTACATGCCTGTTCGCGGTACG 1461

Qy    410 IleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle 429
Db   1462 GTTTTGCTGGATGGCCATGACATCAAGGACCTCAATCTCCGCTGGCTTCGCCAACAGATC 1521

Qy    430 GlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr 449
Db   1522 TCTTTGGTTAGCCAGGAGCCTGTTCTTTTTGGCACGACGATTTATAAGAATATTAGGCAC 1581

Qy    450 Gly--------------------------ArgAspAspValThrAspGluGluMetGlu 460
Db   1582 GGTCTCATCGGCACAAAGTACGAGAATGAATCCGAGGATAAGGTCCGGGAACTCATCGAG 1641

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   1642 AACGCGGCAAAAATGGCGAATGCTCATGACTTTATTACTGCCTTGCCTGAAGGTTATGAG 1701

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   1702 ACCAATGTTGGGCAGCGTGGCTTTCTCCTTTCAGGTGGCCAGAAACAGCGCATTGCAATC 1761

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   1762 GCCCGTGCCGTTGTTAGTGACCCAAAAATCCTGCTCCTGGATGAAGCTACTTCGGCCTTG 1821

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db   1822 GACACAAAATCCGAAGGCGTGGTTCAAGCAGCTTTGGAGAGGGCAGCTGAAGGCCGAACT 1881

Qy    541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db   1882 ACTATTGTGATCGCTCATCGCCTTTCCACGATCAAAACGGCGCACAACATTGTGGTTCTG 1941

Qy    561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db   1942 GTCAATGGCAAAATTGCTGAACAAGGAACTCACGATGAATTGGTTGACCGCGGAGGCGCT 2001

Qy    581 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet 600
Db   2002 TATCGCAAACTTGTGGAGGCTCAACGTATCAATGAACAGAAGGAAGCTGACGCCTTGGAG 2061

Qy    601 ThrTyrSerThrGluArgLysThrAsnSer------LeuProLeuHisSerValLysSer 618
Db   2062 GACGCCGACGCTGAGGATCTCACGAATGCAGATATTGCCAAAATCAAAACTGCGTCAAGC 2121

Qy    619 IleLysSerAspPhe-----------------IleAspLys------AlaGluGluSer 630
Db   2122 GCATCATCCGATCTCGACGGAAAACCCACAACCATTGACCGCACGGGCACCCACAAGTCT 2181

Qy    631 ThrGlnSerLysGluIleSer-----------------LeuProGluValSerLeuLeu 644
Db   2182 GTTTCCAGCGCGATTCTTTCTAAAAGACCCCCCGAAACAACTCCGAAATACTCATTATGG 2241

Qy    645 LysIleLeuLys-----------LeuAsnLysProGluTrpProPheValValLeuGly 660
Db   2242 ACGCTGCTCAAATTTGTTGCTTCCTTCAACCGCCCTGAAATCCCGTACATGCTCATCGGT 2301

Qy    661 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 680
Db   2302 CTTGTCTTCTCAGTGTTAGCTGGTGGTGGCCAACCCACGCAAGCAGTGCTATATGCTAAA 2361

Qy    681 IleIleThrMetPheGly------AsnAsnAspLysThrThrLeuLysHisAspAlaGlu 698
Db   2362 GCCATCAGCACACTCTCGCTCCCAGAATCACAATATAGCAAGCTTCGACATGATGCGGAT 2421

Qy    699 IleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGln 718
Db   2422 TTCTGGTCATTGATGTTCTTCGTGGTTGGTATCATTCAGTTTATCACGCAGTCAACCAAT 2481

Qy    719 GlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPhe 738
Db   2482 GGTGCTGCATTTGCCGTATGCTCCGAGAGACTTATTCGTCGCGCGAGAAGCACTGCCTTT 2541

Qy    739 LysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGly 758
Db   2542 CGGACGATACTCCGTCAAGACATTGCTTTCTTTGACAAGGAAGAGAATAGCACCGGCGCT 2601

Qy    759 LeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIle 778
Db   2602 CTGACCTCTTTCCTGTCCACCGAGACGAAGCATCTCTCCGGTGTTAGCGGTGTGACTCTA 2661

Qy    779 GlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyr 798
Db   2662 GGCACGATCTTGATGACCTCCACGACCCTAGGAGCGGCTATCATTATTGCCCTGGCGATT 2721

Qy    799 GlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMet 818
Db   2722 GGGTGGAAATTGGCCTTAGTTTGTATCTCGGTTGTGCCGGTTCTCCTGGCATGCGGTTTC 2781

Qy    819 IleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAla 838
Db   2782 TACCGATTCTATATGCTAGCCCAGTTTCAATCACGCTCCAAGCTTGCTTATGAGGGATCT 2841

Qy    839 GlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGlu 858
Db   2842 GCAAACTTTGCTTGCGAGGCTACATCGTCTATCCGCACAGTTGCGTCATTAACCCGGGAA 2901

Qy    859 LysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLys 878
Db   2902 AGGGATGTCTGGGAGATTTACCATGCCCAGCTTGACGCACAAGGCAGGACCAGTCTAATC 2961

Qy    879 LysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyr 898
Db   2962 TCTGTCTTGAGGTCATCCCTGTTATATGCGTCGTCGCAGGCACTTGTTTTCTTCTGCGTT 3021

Qy    899 AlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGly 918
Db   3022 GCGCTCGGGTTTTGGTACGGAGGGACACTTCTTGGTCACCACGAGTATGACATTTTCCGC 3081

Qy    919 MetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuVal 938
Db   3082 TTCTTTGTTTGTTTCTCCGAGATTCTCTTTGGTGCTCAATCCGCGGGCACCGTCTTTTCC 3141

Qy    939 LeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGlu 958
Db   3142 TTTGCACCAGACATGGGCAAGGCGAAGAATGCGGCCGCCGAATTCCGACGACTGTTCGAC 3201

Qy    959 LysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGly 978
```

```
Db    3202 CGAAAGCCACAAATTGATAACTGGTCTGAAGAGGGCGAGAAGCTCGAAACGGTGGAAGGT 3261

Qy     979 AsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeu 998
              :::||||||||   |||   |||   ||||||  ||||||:::      :::|||
Db    3262 GAAATCGAATTTAGGAACGTGCACTTCAGATACCCGACCCGCCCAGAACAGCCTGTCCTG 3321

Qy     999 ArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGly 1018
           |||||||||   |||:::::::::  |||:::   ||||||   ||||||   ||||||
Db    3322 CGCGGCTTGGACCTGACCGTGAAGCCTGGACAATATGTTGCGCTTGTCGGACCCAGCGGT 3381

Qy    1019 CysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGln 1038
           |||||||||||||||::::::   ||||||:::|||   ||||||   :::    |||
Db    3382 TGTGGCAAGAGTACCACCATTGCATTGCTTGAGCGCTTTTACGATGCGATTGCCGGGTCC 3441

Qy    1039 ValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIle 1058
           :::|||   ||||||   |||       :::||||||:::         ||||||   :::
Db    3442 ATCCTTGTTGATGGGAAGGACATAAGTAAACTAAATATCAACTCCTACCGCAGCTTTCTG 3501

Qy    1059 AlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyr 1078
           ::::::|||   |||||||||   |||:::      :::|||   |||||||||
Db    3502 TCACTGGTCAGCCAGGAGCCGACACTGTACCAGGGCACCATCAAGGAAAACATCTTACTT 3561

Qy    1079 GlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsn 1098
           |||               ||||||   :::   :::   :::|||         ||||||
Db    3562 GGTATTGTCGAAGATGACGTACCGGAAGAATTCTTGATTAAGGCTTGCAAGGACGCTAAT 3621

Qy    1099 IleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGly 1118
           |||:::   ||||||   |||||||||      :::||||||   ||||||   ||||||
Db    3622 ATCTACGACTTCATCATGTCGCTCCCGGAGGGCTTTAATACAGTTGTTGGCAGCAAGGGA 3681

Qy    1119 AlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLys 1138
                    |||||||||||||||||||:::||||||||||||||||||||||||:::
Db    3682 GGCATGTTGTCTGGCGGCCAAAAGCAACGTGTGGCCATTGCCCGAGCCCTTCTTCGGGAT 3741

Qy    1139 ProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysVal 1158
           ||||||||||||||||||||||||||||||||||||||||||::::::|||||||||||
Db    3742 CCCAAAATCCTTCTTCTCGATGAAGCGACGTCAGCCCTCGACTCCGAGTCAGAAAAGGTC 3801

Qy    1159 ValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArg 1178
           ||||||   |||||||||   |||      |||||||||   :::   |||   ||||||
Db    3802 GTCCAGGCGGCTTTGGATGCCGCTGCCCGAGGCCGAACCACAATCGCCGTTGCACACCGA 3861

Qy    1179 LeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGlu 1198
           ||||||   ||||||   ||||||:::|||   |||         |||||||||   |||
Db    3862 CTCAGCACGATTCAAAAGGCGGACGTTATCTATGTTTTCGACCAAGGCAAGATCGTCGAA 3921

Qy    1199 GlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 1218
                 |||||||||   ||||||::::::   :::      |||::::::|||||||||
Db    3922 AGCGGAACGCACAGCGAACTGGTCCAGAAAAAGGGCCGGTACTACGAGCTGGTCAACTTG 3981

Qy    1219 GlnSerVal 1221
           ||||||:::
Db    3982 CAGAGCTTG 3990
```

RESULT 14

```
US-08-996-545-3
; Sequence 3, Application US/08996545
; Patent No. 5928898
;  GENERAL INFORMATION:
;    APPLICANT:  Skatrud, Paul L.
;    APPLICANT:  de Waard, Maarten A.
;    APPLICANT:  Peery, Robert B.
;    APPLICANT:  Andrade, Alan C.
;    TITLE OF INVENTION:  Multiple Drug Resistance Gene atrD of
;    TITLE OF INVENTION:  Aspergillus Nidulans
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  U.S.
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/996,545
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Webster, Thomas D.
;      REGISTRATION NUMBER:  39,872
;      REFERENCE/DOCKET NUMBER:  X-11766
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-3334
;      TELEFAX:  317-276-2763
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4002 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  mRNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
US-08-996-545-3

Alignment Scores:
Pred. No.:              6.01e-243     Length:          4002
Score:                  2259.00       Matches:         494
Percent Similarity:     58.41%        Conservative:    232
Best Local Similarity:  39.74%        Mismatches:      446
Query Match:            36.68%        Indels:          72
DB:                     2             Gaps:            12

US-09-873-409-5 (1-1222) x US-08-996-545-3 (1-4002)

Qy      15 LeuProLeuMetProLeuValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu 34
           :::||||||:::|||         |||                           |||
Db     368 GUGCCAUUGCUGCCGCGUCGACUUUCCAGA------------------GGAUAAUGUUAU 409

Qy      35 ValGlnThrAsnThrTyrSerPhePhe------------ArgLeuThrLeuTyrTyrVal 50
           :::::::::   |||      |||:::              :::      ||||||:::|||
Db     410 AUCAAAUCUCGUACGAC-GAGUUCUAUGAUGAAUUGACCAAGAACGUACUGUACUUCGUA 468

Qy      51 GlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThr 70
              :::|||:::         :::         |||:::               :::|||   |||
Db     469 UACCUCGGUAUCGGCGAGUUUGUCACUGUCUAUGUUAGUACUGUUGGCUUCAUCUAUACC 528

Qy      71 AlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAsp 90
                    |||::::::||||||:::   :::         |||:::|||   |||:::
Db     529 GGAGAACACGCCACGCAGAAGAUCCGCGAGUAUUACCUUGAGUCUAUCCUGCGCCAGAAC 588

Qy      91 IleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMetThr---AspIle 109
           ||||||:::||||||         ||||||:::   ||||||:::|||   |||
Db     589 AUUGGCUAUUUUGAUAAACUCGGUGCCGGGGAAGUGACCACCCGUAUAACAGCCGAUACA 648

Qy     110 AspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThr 129
           :::   |||   |||||||||   :::|||:::   |||            ::::::|||
Db     649 AACCUUAUCCAGGAUGGCAUUUCGGAGAAGGUCGGUCUCACUUUGACUGCCCUGGCGACA 708

Qy     130 PheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeu 149
           |||               :::      ||||||   |||||||||   |||:::
Db     709 UUCGUGACAGCAUUCAUUAUCGCCUACGUCAAAUACUGGAAGUUGGCUCUAAUUUGCAGC 768

Qy     150 SerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeu 169
           ||||||         ::::::   :::               |||:::   :::|||
Db     769 UCAACAAUUGUGGCCCUCGUUCUCACCAUGGGCGGUGGUUCUCAGUUUAUCAUCAAGUAC 828

Qy     170 ThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSer 189
           :::   |||   |||   :::|||         |||   |||||||||||||||:::|||
Db     829 AGCAAAAAGUCGCUUGACAGCUACGGUGCAGGCGGCACUGUUGCGGAAGAGGUCAUCAGC 888
```

```
Qy    190 SerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeu 209
Db    889 UCCAUCAGAAAUGCCACAGCGUUUGGCACCCAAGACAAGCUUGCGAAG------------ 936

Qy    210 LeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu 229
Db    936 ------------------------------------------------------------ 936

Qy    230 ***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThr 249
Db    937 ---------CAGUAUGAGGUCCACUUAGACGAAGCUGAGAAAUGGGGAACAAAGAACCAG 987

Qy    250 IleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeu 269
Db    988 AUUGUCAUGGGUUUCAUGAUUGGCGCCAUGUUUGGCCUUAUGUACUCGAACUACGGUCUU 1047

Qy    270 AlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThr 289
Db   1048 GGCUUCUGGAUGGGUUCUCGUUUCCUGGUAGAUGGU------GCAGUCGAUGUGGGUGAU 1101

Qy    290 ValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValPro 309
Db   1102 AUUCUCACAGUUCUCAUGGCCAUCUUGAUCGGAUCGUUCUCCUUGGGGAACGUUAGUCCA 1161

Qy    310 HisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAsp 329
Db   1162 AAUGCUCAAGCAUUUACAAACGCUGUGGCCGCGGCCGCAAAGAUAUUUGGAACGAUCGAU 1221

Qy    330 LysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGly 349
Db   1222 CGCCAGUCCCCAUUAGAUCCAUAUUCGAACGAAGGGAAGACGCUCGACCAUUUUGAGGGC 1281

Qy    350 ThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeu 369
Db   1282 CACAUUGAGUUACGCAAUGUCAAGCAUAUUUACCCAUCUAGACCCGAGGUCACCGUCAUG 1341

Qy    370 LysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGly 389
Db   1342 GAGGAUGUUUCUCUGUCAAUGCCCGCUGGAAAAACAACCGCUUUAGUCGGCCCCUCUGGC 1401

Qy    390 SerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPhe 409
Db   1402 UCUGGAAAAAGUACGGUGGUCGGCUUGGUUGAGCGAUUCUACAUGCCUGUUCGCGGUACG 1461

Qy    410 IleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle 429
Db   1462 GUUUUGCUGGAUGGCCAUGACAUCAAGGACCUCAAUCUCCGCUGGCUUCGCCAACAGAUC 1521

Qy    430 GlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr 449
Db   1522 UCUUUGGUUAGCCAGGAGCCUGUUCUUUUUGGCACGACGAUUUAUAAGAAUAUUAGGCAC 1581

Qy    450 Gly---------------------------ArgAspAspValThrAspGluGluMetGlu 460
Db   1582 GGUCUCAUCGGCACAAAGUACGAGAAUGAAUCCGAGGAUAAGGUCCGGGAACUCAUCGAG 1641

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   1642 AACGCGGCAAAAAUGGCGAAUGCUCAUGACUUUAUUACUGCCUUGCCUGAAGGUUAUGAG 1701

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   1702 ACCAAUGUUGGGCAGCGUGGCUUUCUCCUUUCAGGUGGCCAGAAACAGCGCAUUGCAAUC 1761

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   1762 GCCCGUGCCGUUGUUAGUGACCCAAAAAUCCUGCUCCUGGAUGAAGCUACUUCGGCCUUG 1821

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db   1822 GACACAAAAUCCGAAGGCGUGGUUCAAGCAGCUUUGGAGAGGGCAGCUGAAGGCCGAACU 1881

Qy    541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db   1882 ACUAUUGUGAUCGCUCAUCGCCUUUCCACGAUCAAAACGGCGCACAACAUUGUGGUUCUG 1941

Qy    561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db   1942 GUCAAUGGCAAAAUUGCUGAACAAGGAACUCACGAUGAAUUGGUUGACCGCGGAGGCGCU 2001

Qy    581 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet 600
Db   2002 UAUCGCAAACUUGUGGAGGCUCAACGUAUCAAUGAACAGAAGGAAGCUGACGCCUUGGAG 2061

Qy    601 ThrTyrSerThrGluArgLysThrAsnSer------LeuProLeuHisSerValLysSer 618
Db   2062 GACGCCGACGCUGAGGAUCUCACGAAUGCAGAUAUUGCCAAAAUCAAAACUGCGUCAAGC 2121

Qy    619 IleLysSerAspPhe------------------IleAspLys------AlaGluGluSer 630
Db   2122 GCAUCAUCCGAUCUCGACGGAAAACCCACAACCAUUGACCGCACGGGCACCCACAAGUCU 2181

Qy    631 ThrGlnSerLysGluIleSer------------------LeuProGluValSerLeuLeu 644
Db   2182 GUUUCCAGCGCGAUUCUUUCUAAAAGACCCCCCGAAACAACUCCGAAAUACUCAUUAUGG 2241

Qy    645 LysIleLeuLys------------LeuAsnLysProGluTrpProPheValValLeuGly 660
Db   2242 ACGCUGCUCAAAUUUGUUGCUUCCUUCAACCGCCCUGAAAUCCCGUACAUGCUCAUCGGU 2301

Qy    661 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 680
Db   2302 CUUGUCUUCUCAGUGUUAGCUGGUGGUGGCCAACCCACGCAAGCAGUGCUAUAUGCUAAA 2361

Qy    681 IleIleThrMetPheGly------AsnAsnAspLysThrThrLeuLysHisAspAlaGlu 698
Db   2362 GCCAUCAGCACACUCUCGCUCCCAGAAUCACAAUAUAGCAAGCUUCGACAUGAUGCGGAU 2421

Qy    699 IleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGln 718
Db   2422 UUCUGGUCAUUGAUGUUCUUCGUGGUUGGUAUCAUUCAGUUUAUCACGCAGUCAACCAAU 2481

Qy    719 GlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPhe 738
Db   2482 GGUGCUGCAUUUGCCGUAUGCUCCGAGAGACUUAUUCGUCGCGCGAGAAGCACUGCCUUU 2541

Qy    739 LysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGly 758
Db   2542 CGGACGAUACUCCGUCAAGACAUUGCUUUCUUUGACAAGGAAGAGAAUAGCACCGGCGCU 2601

Qy    759 LeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIle 778
Db   2602 CUGACCUCUUUCCUGUCCACCGAGACGAAGCAUCUCUCCGGUGUUAGCGGUGUGACUCUA 2661

Qy    779 GlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyr 798
Db   2662 GGCACGAUCUUGAUGACCUCCACGACCCUAGGAGCGGCUAUCAUUAUUGCCCUGGCGAUU 2721

Qy    799 GlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMet 818
Db   2722 GGGUGGAAAUUGGCCUUAGUUUGUAUCUCGGUUGUGCCGGUUCUCCUGGCAUGCGGUUUC 2781

Qy    819 IleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAla 838
Db   2782 UACCGAUUCUAUAUGCUAGCCCAGUUUCAAUCACGCUCCAAGCUUGCUUAUGAGGGAUCU 2841

Qy    839 GlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGlu 858
Db   2842 GCAAACUUUGCUUGCGAGGCUACAUCGUCUAUCCGCACAGUUGCGUCAUUAACCCGGGAA 2901

Qy    859 LysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLys 878
Db   2902 AGGGAUGUCUGGGAGAUUUACCAUGCCCAGCUUGACGCACAAGGCAGGACCAGUCUAAUC 2961

Qy    879 LysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyr 898
```

```
Db   2962 UCUGUCUUGAGGUCAUCCCUGUUAUAUGCGUCGUCGCAGGCACUUGUUUUCUUCUGCGUU 3021

Qy    899 AlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGly 918
          |||   ||||||   :::|||       |||:::
Db   3022 GCGCUCGGGUUUUGGUACGGAGGGACACUUCUUGGUCACCACGAGUAUGACAUUUUCCGC 3081

Qy    919 MetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuVal 938
             |||:::   |||:::   |||   :::||||||   :::   |||
Db   3082 UUCUUUGUUUGUUUCUCCGAGAUUCUCUUUGGUGCUCAAUCCGCGGGCACCGUCUUUUCC 3141

Qy    939 LeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGlu 958
             ||||||:::      |||||||||:::   ||||||             |||   :::
Db   3142 UUUGCACCAGACAUGGGCAAGGCGAAGAAUGCGGCCGCCGAAUUCCGACGACUGUUCGAC 3201

Qy    959 LysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGly 978
          :::||||||   ||||||:::   |||:::||||||:::|||   :::|||   ||||||
Db   3202 CGAAAGCCACAAAUUGAUAACUGGUCUGAAGAGGGCGAGAAGCUCGAAACGGUGGAAGGU 3261

Qy    979 AsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeu 998
             :::|||||||||   |||   |||   ||||||   ||||||:::      :::|||
Db   3262 GAAAUCGAAUUUAGGAACGUGCACUUCAGAUACCCGACCCGCCCAGAACAGCCUGUCCUG 3321

Qy    999 ArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGly 1018
          |||||||||   |||:::::::::   |||:::   ||||||   ||||||   ||||||
Db   3322 CGCGGCUUGGACCUGACCGUGAAGCCUGGACAAUAUGUUGCGCUUGUCGGACCCAGCGGU 3381

Qy   1019 CysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGln 1038
          |||||||||||||||::::::   ||||||:::|||   ||||||.:::    |||
Db   3382 UGUGGCAAGAGUACCACCAUUGCAUUGCUUGAGCGCUUUUACGAUGCGAUUGCCGGGUCC 3441

Qy   1039 ValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIle 1058
          :::|||   ||||||   |||      :::||||||:::         ||||||   :::
Db   3442 AUCCUUGUUGAUGGGAAGGACAUAAGUAAACUAAAUAUCAACUCCUACCGCAGCUUUCUG 3501

Qy   1059 AlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyr 1078
          ::::::|||   |||||||||   |||:::      :::|||   |||||||||
Db   3502 UCACUGGUCAGCCAGGAGCCGACACUGUACCAGGGCACCAUCAAGGAAAACAUCUUACUU 3561

Qy   1079 GlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsn 1098
          |||               ||||||   :::   :::   :::|||         ||||||
Db   3562 GGUAUUGUCGAAGAUGACGUACCGGAAGAAUUCUUGAUUAAGGCUUGCAAGGACGCUAAU 3621

Qy   1099 IleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGly 1118
          |||:::   ||||||      |||||||||   :::||||||   ||||||   ||||||
Db   3622 AUCUACGACUUCAUCAUGUCGCUCCCGGAGGGCUUUAAUACAGUUGUUGGCAGCAAGGGA 3681

Qy   1119 AlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLys 1138
                |||||||||||||||||||||||:::||||||||||||||||||||:::
Db   3682 GGCAUGUUGUCUGGCGGCCAAAAGCAACGUGUGGCCAUUGCCCGAGCCCUUCUUCGGGAU 3741

Qy   1139 ProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysVal 1158
          ||||||||||||||||||||||||||||||||||||||||||::::::||||||||||||
Db   3742 CCCAAAAUCCUUCUUCUCGAUGAAGCGACGUCAGCCCUCGACUCCGAGUCAGAAAAGGUC 3801

Qy   1159 ValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArg 1178
          ||||||   |||||||||   |||      |||||||||   :::   |||   ||||||
Db   3802 GUCCAGGCGGCUUUGGAUGCCGCUGCCCGAGGCCGAACCACAAUCGCCGUUGCACACCGA 3861

Qy   1179 LeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGlu 1198
          ||||||   ||||||   ||||||:::|||   |||         |||||||||   |||
Db   3862 CUCAGCACGAUUCAAAAGGCGGACGUUAUCUAUGUUUUCGACCAAGGCAAGAUCGUCGAA 3921

Qy   1199 GlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 1218
                |||||||||   ||||||::::::   :::      |||::::::|||||||||
Db   3922 AGCGGAACGCACAGCGAACUGGUCCAGAAAAAGGGCCGGUACUACGAGCUGGUCAACUUG 3981

Qy   1219 GlnSerVal 1221
          ||||||:::
Db   3982 CAGAGCUUG 3990
```

RESULT 15
US-09-328-320-1
; Sequence 1, Application US/09328320

```
; Patent No. 6228615
;  GENERAL INFORMATION:
;    APPLICANT:  Skatrud, Paul L.
;    APPLICANT:  de Waard, Maarten A.
;    APPLICANT:  Peery, Robert B.
;    APPLICANT:  Andrade, Alan C.
;    TITLE OF INVENTION:  Multiple Drug Resistance Gene atrD of
;    TITLE OF INVENTION:  Aspergillus Nidulans
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  U.S.
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/328,320
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/996,545
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Webster, Thomas D.
;      REGISTRATION NUMBER:  39,872
;      REFERENCE/DOCKET NUMBER:  X-11766
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-3334
;      TELEFAX:  317-276-2763
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4002 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..4002
US-09-328-320-1
```

Alignment Scores:

Pred. No.:	6.01e-243	Length:	4002
Score:	2259.00	Matches:	494
Percent Similarity:	58.41%	Conservative:	232
Best Local Similarity:	39.74%	Mismatches:	446
Query Match:	36.68%	Indels:	72
DB:	4	Gaps:	12

US-09-873-409-5 (1-1222) x US-09-328-320-1 (1-4002)

```
Qy     15 LeuProLeuMetProLeuValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu 34
          :::||||||:::|||      |||                               |||
Db    368 GTGCCATTGCTGCCGCGTCGACTTTCCAGA------------------GGATAATGTTAT 409

Qy     35 ValGlnThrAsnThrTyrSerPhePhe------------ArgLeuThrLeuTyrTyrVal 50
          :::::::::   |||      |||:::               :::      ||||||:::|||
Db    410 ATCAAATCTCGTACGAC-GAGTTCTATGATGAATTGACCAAGAACGTACTGTACTTCGTA 468

Qy     51 GlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThr 70
             :::|||:::         :::      |||:::            :::|||   |||
Db    469 TACCTCGGTATCGGCGAGTTTGTCACTGTCTATGTTAGTACTGTTGGCTTCATCTATACC 528

Qy     71 AlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAsp 90
                |||::::::||||||:::   :::      |||:::|||   |||:::
```

```
Db    529 GGAGAACACGCCACGCAGAAGATCCGCGAGTATTACCTTGAGTCTATCCTGCGCCAGAAC 588

Qy     91 IleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMetThr---AspIle 109
Db    589 ATTGGCTATTTTGATAAACTCGGTGCCGGGGAAGTGACCACCCGTATAACAGCCGATACA 648

Qy    110 AspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThr 129
Db    649 AACCTTATCCAGGATGGCATTTCGGAGAAGGTCGGTCTCACTTTGACTGCCCTGGCGACA 708

Qy    130 PheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeu 149
Db    709 TTCGTGACAGCATTCATTATCGCCTACGTCAAATACTGGAAGTTGGCTCTAATTTGCAGC 768

Qy    150 SerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeu 169
Db    769 TCAACAATTGTGGCCCTCGTTCTCACCATGGGCGGTGGTTCTCAGTTTATCATCAAGTAC 828

Qy    170 ThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSer 189
Db    829 AGCAAAAAGTCGCTTGACAGCTACGGTGCAGGCGGCACTGTTGCGGAAGAGGTCATCAGC 888

Qy    190 SerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeu 209
Db    889 TCCATCAGAAATGCCACAGCGTTTGGCACCCAAGACAAGCTTGCGAAG------------ 936

Qy    210 LeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu 229
Db    936 ------------------------------------------------------------ 936

Qy    230 ***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThr 249
Db    937 ---------CAGTATGAGGTCCACTTAGACGAAGCTGAGAAATGGGGAACAAAGAACCAG 987

Qy    250 IleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeu 269
Db    988 ATTGTCATGGGTTTCATGATTGGCGCCATGTTTGGCCTTATGTACTCGAACTACGGTCTT 1047

Qy    270 AlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThr 289
Db   1048 GGCTTCTGGATGGGTTCTCGTTTCCTGGTAGATGGT------GCAGTCGATGTGGGTGAT 1101

Qy    290 ValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValPro 309
Db   1102 ATTCTCACAGTTCTCATGGCCATCTTGATCGGATCGTTCTCCTTGGGGAACGTTAGTCCA 1161

Qy    310 HisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAsp 329
Db   1162 AATGCTCAAGCATTTACAAACGCTGTGGCCGCGGCCGCAAAGATATTTGGAACGATCGAT 1221

Qy    330 LysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGly 349
Db   1222 CGCCAGTCCCCATTAGATCCATATTCGAACGAAGGGAAGACGCTCGACCATTTTGAGGGC 1281

Qy    350 ThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeu 369
Db   1282 CACATTGAGTTACGCAATGTCAAGCATATTTACCCATCTAGACCCGAGGTCACCGTCATG 1341

Qy    370 LysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGly 389
Db   1342 GAGGATGTTTCTCTGTCAATGCCCGCTGGAAAAACAACCGCTTTAGTCGGCCCCTCTGGC 1401

Qy    390 SerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPhe 409
Db   1402 TCTGGAAAAAGTACGGTGGTCGGCTTGGTTGAGCGATTCTACATGCCTGTTCGCGGTACG 1461

Qy    410 IleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle 429
Db   1462 GTTTTGCTGGATGGCCATGACATCAAGGACCTCAATCTCCGCTGGCTTCGCCAACAGATC 1521

Qy    430 GlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr 449
Db   1522 TCTTTGGTTAGCCAGGAGCCTGTTCTTTTTGGCACGACGATTTATAAGAATATTAGGCAC 1581

Qy    450 Gly--------------------------ArgAspAspValThrAspGluGluMetGlu 460
Db   1582 GGTCTCATCGGCACAAAGTACGAGAATGAATCCGAGGATAAGGTCCGGGAACTCATCGAG 1641

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   1642 AACGCGGCAAAAATGGCGAATGCTCATGACTTTATTACTGCCTTGCCTGAAGGTTATGAG 1701

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   1702 ACCAATGTTGGGCAGCGTGGCTTTCTCCTTTCAGGTGGCCAGAAACAGCGCATTGCAATC 1761

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   1762 GCCCGTGCCGTTGTTAGTGACCCAAAAATCCTGCTCCTGGATGAAGCTACTTCGGCCTTG 1821

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db   1822 GACACAAAATCCGAAGGCGTGGTTCAAGCAGCTTTGGAGAGGGCAGCTGAAGGCCGAACT 1881

Qy    541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db   1882 ACTATTGTGATCGCTCATCGCCTTTCCACGATCAAAACGGCGCACAACATTGTGGTTCTG 1941

Qy    561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db   1942 GTCAATGGCAAAATTGCTGAACAAGGAACTCACGATGAATTGGTTGACCGCGGAGGCGCT 2001

Qy    581 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet 600
Db   2002 TATCGCAAACTTGTGGAGGCTCAACGTATCAATGAACAGAAGGAAGCTGACGCCTTGGAG 2061

Qy    601 ThrTyrSerThrGluArgLysThrAsnSer------LeuProLeuHisSerValLysSer 618
Db   2062 GACGCCGACGCTGAGGATCTCACGAATGCAGATATTGCCAAAATCAAAACTGCGTCAAGC 2121

Qy    619 IleLysSerAspPhe-----------------IleAspLys------AlaGluGluSer 630
Db   2122 GCATCATCCGATCTCGACGGAAAACCCACAACCATTGACCGCACGGGCACCCACAAGTCT 2181

Qy    631 ThrGlnSerLysGluIleSer-----------------LeuProGluValSerLeuLeu 644
Db   2182 GTTTCCAGCGCGATTCTTTCTAAAAGACCCCCCGAAACAACTCCGAAATACTCATTATGG 2241

Qy    645 LysIleLeuLys-----------LeuAsnLysProGluTrpProPheValValLeuGly 660
Db   2242 ACGCTGCTCAAATTTGTTGCTTCCTTCAACCGCCCTGAAATCCCGTACATGCTCATCGGT 2301

Qy    661 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 680
Db   2302 CTTGTCTTCTCAGTGTTAGCTGGTGGTGGCCAACCCACGCAAGCAGTGCTATATGCTAAA 2361

Qy    681 IleIleThrMetPheGly------AsnAsnAspLysThrThrLeuLysHisAspAlaGlu 698
Db   2362 GCCATCAGCACACTCTCGCTCCCAGAATCACAATATAGCAAGCTTCGACATGATGCGGAT 2421

Qy    699 IleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGln 718
Db   2422 TTCTGGTCATTGATGTTCTTCGTGGTTGGTATCATTCAGTTTATCACGCAGTCAACCAAT 2481

Qy    719 GlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPhe 738
Db   2482 GGTGCTGCATTTGCCGTATGCTCCGAGAGACTTATTCGTCGCGCGAGAAGCACTGCCTTT 2541

Qy    739 LysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGly 758
Db   2542 CGGACGATACTCCGTCAAGACATTGCTTTCTTTGACAAGGAAGAGAATAGCACCGGCGCT 2601

Qy    759 LeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIle 778
Db   2602 CTGACCTCTTTCCTGTCCACCGAGACGAAGCATCTCTCCGGTGTTAGCGGTGTGACTCTA 2661
```

```
Qy    779 GlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyr 798
Db   2662 GGCACGATCTTGATGACCTCCACGACCCTAGGAGCGGCTATCATTATTGCCCTGGCGATT 2721

Qy    799 GlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMet 818
Db   2722 GGGTGGAAATTGGCCTTAGTTTGTATCTCGGTTGTGCCGGTTCTCCTGGCATGCGGTTTC 2781

Qy    819 IleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAla 838
Db   2782 TACCGATTCTATATGCTAGCCCAGTTTCAATCACGCTCCAAGCTTGCTTATGAGGGATCT 2841

Qy    839 GlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGlu 858
Db   2842 GCAAACTTTGCTTGCGAGGCTACATCGTCTATCCGCACAGTTGCGTCATTAACCCGGGAA 2901

Qy    859 LysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLys 878
Db   2902 AGGGATGTCTGGGAGATTTACCATGCCCAGCTTGACGCACAAGGCAGGACCAGTCTAATC 2961

Qy    879 LysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyr 898
Db   2962 TCTGTCTTGAGGTCATCCCTGTTATATGCGTCGTCGCAGGCACTTGTTTTCTTCTGCGTT 3021

Qy    899 AlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGly 918
Db   3022 GCGCTCGGGTTTTGGTACGGAGGGACACTTCTTGGTCACCACGAGTATGACATTTTCCGC 3081

Qy    919 MetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuVal 938
Db   3082 TTCTTTGTTTGTTTCTCCGAGATTCTCTTTGGTGCTCAATCCGCGGGCACCGTCTTTTCC 3141

Qy    939 LeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGlu 958
Db   3142 TTTGCACCAGACATGGGCAAGGCGAAGAATGCGGCCGCCGAATTCCGACGACTGTTCGAC 3201

Qy    959 LysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGly 978
Db   3202 CGAAAGCCACAAATTGATAACTGGTCTGAAGAGGGCGAGAAGCTCGAAACGGTGGAAGGT 3261

Qy    979 AsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeu 998
Db   3262 GAAATCGAATTTAGGAACGTGCACTTCAGATACCCGACCCGCCCAGAACAGCCTGTCCTG 3321

Qy    999 ArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGly 1018
Db   3322 CGCGGCTTGGACCTGACCGTGAAGCCTGGACAATATGTTGCGCTTGTCGGACCCAGCGGT 3381

Qy   1019 CysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGln 1038
Db   3382 TGTGGCAAGAGTACCACCATTGCATTGCTTGAGCGCTTTTACGATGCGATTGCCGGGTCC 3441

Qy   1039 ValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIle 1058
Db   3442 ATCCTTGTTGATGGGAAGGACATAAGTAAACTAAATATCAACTCCTACCGCAGCTTTCTG 3501

Qy   1059 AlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyr 1078
Db   3502 TCACTGGTCAGCCAGGAGCCGACACTGTACCAGGGCACCATCAAGGAAAACATCTTACTT 3561

Qy   1079 GlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsn 1098
Db   3562 GGTATTGTCGAAGATGACGTACCGGAAGAATTCTTGATTAAGGCTTGCAAGGACGCTAAT 3621

Qy   1099 IleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGly 1118
Db   3622 ATCTACGACTTCATCATGTCGCTCCCGGAGGGCTTTAATACAGTTGTTGGCAGCAAGGGA 3681

Qy   1119 AlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLys 1138
Db   3682 GGCATGTTGTCTGGCGGCCAAAAGCAACGTGTGGCCATTGCCCGAGCCCTTCTTCGGGAT 3741

Qy   1139 ProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysVal 1158
Db   3742 CCCAAAATCCTTCTTCTCGATGAAGCGACGTCAGCCCTCGACTCCGAGTCAGAAAAGGTC 3801

Qy   1159 ValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArg 1178
Db   3802 GTCCAGGCGGCTTTGGATGCCGCTGCCCGAGGCCGAACCACAATCGCCGTTGCACACCGA 3861

Qy   1179 LeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGlu 1198
Db   3862 CTCAGCACGATTCAAAAGGCGGACGTTATCTATGTTTTCGACCAAGGCAAGATCGTCGAA 3921

Qy   1199 GlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 1218
Db   3922 AGCGGAACGCACAGCGAACTGGTCCAGAAAAAGGGCCGGTACTACGAGCTGGTCAACTTG 3981

Qy   1219 GlnSerVal 1221
Db   3982 CAGAGCTTG 3990
```

Search completed: March 31, 2003, 15:32:18
Job time : 455.956 secs

```
OM protein - nucleic search, using frame_plus_p2n model

Run on:          March 30, 2003, 03:22:44 ; Search time 8405.14 Seconds
                                            (without alignments)
                                            2354.618 Million cell updates/sec

Title:           US-09-873-409-5
Perfect score:   6159
Sequence:        1 MILGILASLVNGACLPLMPL..........QELLRNRDIYFKLVNAQSVQ 1222

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters:      32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115420_19240/app_query.fasta_1.7544
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_30544_@runat_27032003_115420_19240 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       EST:*
                 1:  em_estba:*
                 2:  em_esthum:*
                 3:  em_estin:*
                 4:  em_estmu:*
                 5:  em_estov:*
                 6:  em_estpl:*
                 7:  em_estro:*
                 8:  em_htc:*
                 9:  gb_est1:*
                 10:  gb_est2:*
                 11:  gb_htc:*
                 12:  gb_est3:*
                 13:  gb_est4:*
                 14:  gb_est5:*
                 15:  em_estfun:*
                 16:  em_estom:*
                 17:  gb_gss:*
                 18:  em_gss_hum:*
                 19:  em_gss_inv:*
                 20:  em_gss_pln:*
                 21:  em_gss_vrt:*
                 22:  em_gss_fun:*
                 23:  em_gss_mam:*
                 24:  em_gss_mus:*
                 25:  em_gss_other:*
                 26:  em_gss_pro:*
                 27:  em_gss_rod:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1092	17.7	669	13	BM013981	BM013981 603639619
	2	1051	17.1	1033	14	BM926413	BM926413 AGENCOURT
	3	1008	16.4	943	9	AL520322	AL520322 AL520322
	4	995	16.2	760	13	BM016204	BM016204 603642659
	5	918	14.9	545	9	AL040762	AL040762 DKFZp434C
	6	903	14.7	929	14	BQ882401	BQ882401 AGENCOURT
	7	856.5	13.9	948	13	BM471690	BM471690 AGENCOURT
	8	834	13.5	998	14	BM904842	BM904842 AGENCOURT
	9	833.5	13.5	2676	11	AK014319	AK014319 Mus muscu
	10	771	12.5	894	12	BF584668	BF584668 602098406
	11	770.5	12.5	1019	12	BG248052	BG248052 602359987
	12	753	12.2	726	12	BG293345	BG293345 602390738
	13	752.5	12.2	871	17	AZ682350	AZ682350 ENTKB16TF
	14	751	12.2	944	12	BF796582	BF796582 602258463
	15	744.5	12.1	872	11	AK020318	AK020318 Mus muscu
	16	743.5	12.1	886	17	BH139685	BH139685 ENTNA47TR
c	17	734.5	11.9	932	17	AZ670821	AZ670821 ENTJN69TF
	18	728.5	11.8	947	17	AZ683753	AZ683753 ENTIL96TF
	19	726.5	11.8	886	17	AZ540627	AZ540627 ENTEQ18TF
c	20	723.5	11.7	880	17	AZ687805	AZ687805 ENTLU52TF
	21	721.5	11.7	939	14	BQ720763	BQ720763 AGENCOURT
	22	720.5	11.7	897	17	AZ541090	AZ541090 ENTDS67TR
	23	719	11.7	780	12	BG587938	BG587938 EST489713
	24	711.5	11.6	913	17	BH155700	BH155700 ENTRO54TR
	25	709.5	11.5	834	17	AZ548312	AZ548312 ENTFG07TR
c	26	708.5	11.5	853	17	AZ679807	AZ679807 ENTHI67TR
	27	705.5	11.5	823	17	AZ532602	AZ532602 ENTCR50TR
	28	697.5	11.3	861	14	BQ717101	BQ717101 AGENCOURT
	29	692.5	11.2	1341	11	AY108285	AY108285 Zea mays
	30	690.5	11.2	899	17	BH154857	BH154857 ENTRO47TF
c	31	690	11.2	835	13	BJ349604	BJ349604 BJ349604
c	32	689.5	11.2	891	17	AZ682250	AZ682250 ENTKT68TR
	33	688	11.2	1123	11	AY108485	AY108485 Zea mays
	34	687	11.2	715	14	BQ869512	BQ869512 QGD6G09.y
c	35	684	11.1	820	13	BJ356540	BJ356540 BJ356540
	36	683	11.1	600	13	BM486593	BM486593 pgm2n.pk0
	37	682.5	11.1	919	17	AZ690701	AZ690701 ENTIY68TR
	38	680	11.0	926	14	BQ123477	BQ123477 EST609053
	39	680	11.0	1813	11	U66688	U66688 Homo sapien
	40	677.5	11.0	795	12	BF313560	BF313560 601900192
	41	674.5	11.0	815	13	BI409057	BI409057 602961127
	42	670	10.9	852	17	BH720383	BH720383 BOHZZ94TF
	43	669	10.9	547	13	BJ030711	BJ030711 BJ030711
c	44	667	10.8	937	17	BH150760	BH150760 ENTQG42TR
c	45	666	10.8	792	13	BJ352404	BJ352404 BJ352404

ALIGNMENTS

```
RESULT 1
BM013981
LOCUS       BM013981                 669 bp    mRNA    linear   EST 30-OCT-2001
DEFINITION  603639619F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5415556 5',
            mRNA sequence.
ACCESSION   BM013981
VERSION     BM013981.1  GI:16528335
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 669)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
```

```
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12060  row: f  column: 05
            High quality sequence stop: 664.
FEATURES             Location/Qualifiers
     source          1..669
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5415556"
                     /clone_lib="NIH_MGC_87"
                     /tissue_type="mammary adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.383 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      203 a    122 c    162 g    182 t
ORIGIN

Alignment Scores:
Pred. No.:              7.71e-109     Length:         669
Score:                  1092.00       Matches:        215
Percent Similarity:     99.54%        Conservative:   2
Best Local Similarity:  98.62%        Mismatches:     0
Query Match:            17.73%        Indels:         1
DB:                     13            Gaps:           0

US-09-873-409-5 (1-1222) x BM013981 (1-669)

Qy    292 AlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPhe 311
          :::||||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db     17 AGTGTTTTCTTTAATGTAATCCATAGCAGTTATTGCATTGGAGCAGCAGTCCCTCACTTT 76

Qy    312 GluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLys 331
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     77 GAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTTATTGATAAGAAA 136

Qy    332 ProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrVal 351
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    137 CCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTG 196

Qy    352 GluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGly 371
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    197 GAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGT 256

Qy    372 LeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGly 391
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    257 CTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGG 316

Qy    392 LysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMet 411
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    317 AAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATG 376

Qy    412 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 431
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    377 GTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTG 436

Qy    432 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 451
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    437 GTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGA 496

Qy    452 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 471
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    497 GATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTT 556

Qy    472 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 491
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    557 ATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATGAGT 616

Qy    492 GlyGlyGlnLysGlnArg-IleAlaIleAlaArgAlaLeuValArgAsnPro 508
          |||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    617 GGAGGGCAGAAACAGAGGCATCGCAATTGCTCGTGCCTTAGTTCGAAACCCC 668

RESULT 2
BM926413
LOCUS       BM926413                1033 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6600788 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5764845
            5', mRNA sequence.
ACCESSION   BM926413
VERSION     BM926413.1  GI:19376792
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1033)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12818  row: o  column: 22
            High quality sequence start: 3
            High quality sequence stop: 686.
FEATURES             Location/Qualifiers
     source          1..1033
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5764845"
                     /clone_lib="NIH_MGC_114"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
                     male brains, age range 23-27 yo. Library is oligo-dT
                     primed and directionally cloned (EcoRV site is  destroyed
                     upon cloning).  Average insert size 1.5 kb, insert size
                     range 1-3 kb.  Library is normalized and enriched for
                     full-length  clones and was constructed by C. Gruber
                     (Invitrogen).  Research  Genetics tracking code 019. Note:
                     this is a NIH_MGC Library."
BASE COUNT      289 a    212 c    218 g    312 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              5.31e-104     Length:         1033
Score:                  1051.00       Matches:        221
Percent Similarity:     94.49%        Conservative:   2
Best Local Similarity:  93.64%        Mismatches:     13
Query Match:            17.06%        Indels:         3
DB:                     14            Gaps:           0

US-09-873-409-5 (1-1222) x BM926413 (1-1033)

Qy    293 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 312
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 GTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAA 239

Qy    313 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 332
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 ACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTCATTGATAAGAAACCC 299
```

```
Qy     333 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 352
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     300 AGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGAA 359

Qy     353 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 372
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     360 TTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCTG 419

Qy     373 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 392
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     420 AATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAG 479

Qy     393 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 412
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     480 AGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGTG 539

Qy     413 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 432
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     540 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTGGTT 599

Qy     433 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 452
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     600 AGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGAGAT 659

Qy     453 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 472
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     660 GATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATC 719

Qy     473 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 492
           ||||||||||||||   |||||||||||||||||||||   ||||||    |||||||||
Db     720 ATGGAGTTTCCTAATAA-TTTAATACATTGGTAGGGGAANAAGGAGCTCCAATGAGTGGA 778

Qy     493 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 512
             |||     ||||||  ||||||   |||||||||||||||||||:::     |||
Db     779 AGCCAGAACCG-AGGATCCCAATTGCTCCTGCCTTAGTTCGAAACCCCAGGAT-CTGAAT 836

Qy     513 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 528
           ||||||||||||||||||||||||||||||||||:::        |||
Db     837 TTAGATGAAGCTACGTCTGCCCTGGATTCAGAAACCAGTCAACTGGTC 884
```

RESULT 3
AL520322

```
LOCUS       AL520322                 943 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL520322 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB006YC15 5
            prime, mRNA sequence.
ACCESSION   AL520322
VERSION     AL520322.1  GI:12783815
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 943)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .943
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DB006YC15"
                     /clone_lib="LTI_NFL004_NBC2"
                     /sex="male"
                     /tissue_type="neuroblastoma cells"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector. Library was normalized. Library was constructed
                     by Life Technologies. Contact : Feng Liang Life
                     Technologies, a division of Invitrogen 9800 Medical Center
                     Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
                     8371 Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      253 a    225 c    253 g    211 t     1 others
ORIGIN

Alignment Scores:
Pred. No.:              2.36e-99      Length:        943
Score:                  1008.00       Matches:       194
Percent Similarity:     80.40%        Conservative:  48
Best Local Similarity:  64.45%        Mismatches:    58
Query Match:            16.37%        Indels:        1
DB:                     9             Gaps:          0

US-09-873-409-5 (1-1222) x AL520322 (1-943)

Qy     921 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 940
           ::::::|||:::|||:::   :::|||||||||||:::|||:::             |||
Db       5 TTATTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCT 64

Qy     941 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 960
           |||:::|||:::|||||||||      |||||||||:::      ::::::||||||
Db      65 CCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACC 124

Qy     961 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 980
           |||   |||||||||   |||   ||||||      |||:::|||   |||||||||:::
Db     125 CCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTC 184

Qy     981 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 1000
              |||   ||||||   |||   ||||||   |||||||||:::   :::|||:::|||
Db     185 ACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGA 244

Qy    1001 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 1020
           |||||||||   :::::::::|||:::|||:::|||   |||||||||||||||||||||
Db     245 CTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGG 304

Qy    1021 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 1040
           |||||||||   ||||||||||||:::|||   |||||||||:::   |||:::||||||
Db     305 AAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTG 364

Qy    1041 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 1060
              ||||||   :::   |||   |||||||||||||||||||||:::   :::   |||
Db     365 CTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATC 424

Qy    1061 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 1080
           |||   |||||||||:::||||||:::|||||||||||||||||||||||||||||||||
Db     425 GTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGAC 484

Qy    1081 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 1100
           ||||||      |||||||||         ||||||                ||||||||||||
Db     485 AACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACAT 544

Qy    1101 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 1120
           :::|||||||||   ||||||   ||||||:::|||:::||||||   ||||||   |||
Db     545 GCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAG 604

Qy    1121 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 1140
           |||||||||||||||||||||||||||:::||||||||||||||||||:::::::::|||
Db     605 CTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCAT 664

Qy    1141 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 1160
           |||||||||||||||||||||||||||||||||||   :::||||||||||||||||||
Db     665 ATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAA 724

Qy    1161 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 1180
              ||||||||||||||||   ||||||||||||:::|||:::      ||||||||||||
Db     725 GAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCC 784

Qy    1181 AlaIleGlnAsnAlaAspLeuIleVal-ValLeuHisAsnGlyLysIleLysGluGlnGl 1200
```

```
              |||||||||||||||||||||||| |||        ||||||::::::||||||    ||
Db     785 ACCATCCAGAATGCAGACTTAATAGTGTGTTTCAGAATGGCAGAGTCAAGGAGCATGG 844

Qy    1200 yThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSe 1220
            ||||||||||:::||||||         :::  ||||||||||    :::||||:::    |||::
Db     845 CACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGC 904

Qy    1220 r 1220
             :
Db     905 T 905

RESULT 4
BM016204
LOCUS       BM016204                 760 bp    mRNA    linear   EST 30-OCT-2001
DEFINITION  603642659F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5418615 5',
            mRNA sequence.
ACCESSION   BM016204
VERSION     BM016204.1  GI:16530558
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 760)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12068  row: e  column: 16
            High quality sequence stop: 740.
FEATURES             Location/Qualifiers
     source          1..760
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5418615"
                     /clone_lib="NIH_MGC_87"
                     /tissue_type="mammary adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.383 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      229 a    150 c    187 g    193 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              4.27e-98       Length:         760
Score:                  995.00         Matches:        203
Percent Similarity:     98.07%         Conservative:   0
Best Local Similarity:  98.07%         Mismatches:     2
Query Match:            16.16%         Indels:         2
DB:                     13             Gaps:           0

US-09-873-409-5 (1-1222) x BM016204 (1-760)

Qy     330 LysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGly 349
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3 AAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGA 62

Qy     350 ThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeu 369
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63 ACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTG 122

Qy     370 LysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGly 389
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 AAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGC 182

Qy     390 SerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPhe 409
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 AGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTT 242

Qy     410 IleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle 429
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 ATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATN 302

Qy     430 GlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr 449
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 GGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTAT 362

Qy     450 GlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyr 469
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     363 GGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTAT 422

Qy     470 AspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGln 489
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     423 GATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAA 482

Qy     490 MetSerGlyGlyGlnLysGlnArgIleAlaIle-AlaArgAlaLeuValArgAsnProLy 509
              ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db     483 ATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGGCTCGTGCCTTAGTTCGAAACCCCAA 542

Qy     509 sIleLeuIleLeu-AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValG 529
              ||||||||||    ||||||||||||||||||||||||||||||||||||||||||||||
Db     543 GATTCTGATTTTCAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTC 602

Qy     529 lnAlaAlaLeuGluLys 534
              ||||||||||||||||
Db     603 AAGCTGCACTGGAGAAG 619

RESULT 5
AL040762
LOCUS       AL040762                 545 bp    mRNA    linear   EST 29-FEB-2000
DEFINITION  DKFZp434C1815_r1 434 (synonym: htes3) Homo sapiens cDNA clone
            DKFZp434C1815 5', mRNA sequence.
ACCESSION   AL040762
VERSION     AL040762.1  GI:5409708
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 545)
  AUTHORS   Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
            Wiemann,S.
  TITLE     EST (Bloecker, et al.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Bloecker H
            MIPS
            Am Klopferspitz 18a D-82152 Martinsried, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by GBF (National Research Centre for Biotechnology Ltd.,
            Braunschweig/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            s1 sequence also available.
            This clone (DKFZp434C1815) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1..545
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp434C1815"
                     /clone_lib="434 (synonym: htes3)"
                     /tissue_type="testis"
                     /dev_stage="adult"
```

```
                 /lab_host="DH10B"
                 /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT      171 a     89 c     138 g     147 t
ORIGIN

Alignment Scores:
Pred. No.:               6.63e-90       Length:        545
Score:                   918.00         Matches:       181
Percent Similarity:      100.00%        Conservative:  0
Best Local Similarity:   100.00%        Mismatches:    0
Query Match:             14.91%         Indels:        0
DB:                      9              Gaps:          0

US-09-873-409-5 (1-1222) x AL040762 (1-545)

Qy     333 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 352
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3 AGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGAA 62

Qy     353 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 372
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63 TTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCTG 122

Qy     373 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 392
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 AATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAG 182

Qy     393 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 412
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 AGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGTG 242

Qy     413 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 432
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTGGTT 302

Qy     433 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 452
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 AGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGAGAT 362

Qy     453 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 472
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     363 GATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATC 422

Qy     473 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 492
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     423 ATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGA 482

Qy     493 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 512
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     483 GGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATT 542

Qy     513 Leu 513
           |||
Db     543 TTA 545

RESULT 6
BQ882401
LOCUS       BQ882401                 929 bp    mRNA    linear   EST 16-AUG-2002
DEFINITION  AGENCOURT_8627902 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6291782
            5', mRNA sequence.
ACCESSION   BQ882401
VERSION     BQ882401.1  GI:22274409
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 929)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            cDNA Library Preparation: Rubin Laboratory
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2492  row: c  column: 15
            High quality sequence stop: 677.
FEATURES             Location/Qualifiers
     source          1..929
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6291782"
                     /clone_lib="NIH_MGC_43"
                     /tissue_type="normal pigmented retinal epithelium"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Library constructed  by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and  Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library. |"
BASE COUNT      281 a    176 c    237 g    234 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:               7.64e-88       Length:        929
Score:                   903.00         Matches:       188
Percent Similarity:      73.02%         Conservative:  61
Best Local Similarity:   55.13%         Mismatches:    58
Query Match:             14.66%         Indels:        34
DB:                      14             Gaps:          3

US-09-873-409-5 (1-1222) x BQ882401 (1-929)

Qy     176 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThr-ValIl 195
           ||||||:::||||||||||||||||||||||||||||||:::::::||||||||| |||||
Db       1 GCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGGTGAT 60

Qy     195 eAlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTy 215
           |||||||      |||:::||||||||:::
Db      61 TGCATTTGGAGGACAAAAGAAAGAACTTGAA----------------------------- 91

Qy     215 rAlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrTh 235
                                                         ||||||
Db      92 ---------------------------------------------------AGGTACAA 99

Qy     235 rGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSe 255
            :::|||||::::::||||||       ||||||||:::    |||    :::    :::||
Db     100 CAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTC 159

Qy     255 rLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyTh 275
           |:::||||||   :::      :::       :::|||  ||||||||||||||||||||
Db     160 TATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGAC 219

Qy     275 rSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePh 295
           |:::|||:::|||:::||||||       |||:::||||||  ||||||     |||||||
Db     220 CACCTTGGTCCTCTCAGGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTT 273

Qy     295 eSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAl 315
           |||||||:::      ::::::   :::|||   |||   |||        |||   |||||
Db     274 TTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGC 333

Qy     315 aIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAs 335
           |   ||||||||||||||:::     |||||||::::::||||||   |||||||||||||||
Db     334 AAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGA 393

Qy     335 pAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAs 355
           |::::::|||   :::|||:::||||||::::::|||:::|||    :::||||||:::||
Db     394 CAGCTATTCGAAGAGTGGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAA 453
```

```
Qy     355 nValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuAr 375
           ||||   |||:::||||||||||      :::|||||||||||||||||||||||::
Db     454 TGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAGATCTTGAAGGGTCTGAACCTGAA 513

Qy     375 gIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVa 395
           :::::::|||||:::||||||||||||||||    :::|||   |||||||||||
Db     514 GGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAAC 573

Qy     395 lValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAs 415
            ||||||||:::||||||||||||||||    :::|||    :::    |||||
Db     574 AGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACA 633

Qy     415 nAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGl 435
            ||||||||    :::|||||||||        |||:::    ||||||||||||||||||
Db     634 GGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGA 693

Qy     435 uProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValTh 455
           |||||||||||||    ||||||||||:::    ||||||    |||||||||:::::||||
Db     694 ACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCCCTATGGCCGTGAAAATGTCAC 753

Qy     455 rAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla-TyrAspPheIleMetGluP 475
           |    :::|||:::|||:::|||    :::|||||||||||| ||||||||||||||:::
Db     754 CATGGATGAGATTGAGAAAGCTGTCCAGGAAGCCAATGCCCTATGACTTTATCATGAAAC 813

Qy     475 heProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGly---G 494
              |||:::||||||:::|||||||||||    :::        |||:::||||||||    |
Db     814 TGCCTCATAAATTTGACACCCTGGTTGGANAAAGAAGGGGCCAATTGAGTGGTGGGGCAC 873

Qy     494 lnLysGlnArgIleAlaIleAlaArg-AlaLeuVal-ArgAsnProLysIleLeu 511
            ||    :::          :::||||||  |||||||||    ||||||||||||||
Db     874 AAACCAAAGGATCTCCCTTGGCACGTGGCCCTGGTTCCCCAACCCCAAGATCCTC 928
```

RESULT 7
BM471690
LOCUS   BM471690   948 bp   mRNA   linear   EST 05-FEB-2002
DEFINITION AGENCOURT_6465349 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5539117
            5', mRNA sequence.
ACCESSION   BM471690
VERSION   BM471690.1 GI:18520732
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 948)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12233 row: b column: 14
            High quality sequence stop: 569.
FEATURES   Location/Qualifiers
     source   1..948
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5539117"
                     /clone_lib="NIH_MGC_72"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                     Average insert size 2 kb. Library constructed by Life
                     Technologies."
BASE COUNT   285 a   187 c   194 g   282 t

ORIGIN

Alignment Scores:			
Pred. No.:	1e-82	Length:	948
Score:	856.50	Matches:	186
Percent Similarity:	89.67%	Conservative:	5
Best Local Similarity:	87.32%	Mismatches:	14
Query Match:	13.91%	Indels:	8
DB:	13	Gaps:	1

US-09-873-409-5 (1-1222) x BM471690 (1-948)

```
Qy     293 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 312
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     205 GTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAA 264

Qy     313 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 332
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     265 ACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTTATTGATAAGAAACCC 324

Qy     333 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 352
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     325 AGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGAA 384

Qy     353 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 372
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     385 TTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCTG 444

Qy     373 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 392
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     445 AATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAG 504

Qy     393 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 412
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     505 AGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGTG 564

Qy     413 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle-GlyValVa 432
           ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db     565 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTTGGAGTGGT 624

Qy     432 lSerGlnGluPro--ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr-GlyAr 451
           ||||||||||||  ||||||||||||||||||||||||||||||||||||||| |||||
Db     625 TAGTCAAGAGCCCTGGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGGACG 684

Qy     451 gAspAspValThrAspGluGlu-MetGluArgAlaAlaArgGluAla-AsnAlaTyrAsp 470
           ||||||   |||||||||::: ||||||:::|||   ||||||        |||   :::
Db     685 AGATGATGGGACTGATGAACAAAATGGAAAAAGCAACAAGGGAAACCAAAGGCGAATGAA 744

Qy     471 PheIleMetGlu-PheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGl 489
           ||||||||||| ||||||  ||||||||||||||       |||:::|||      ||
Db     745 TTTATCATGGAATTTCCCAAAAAAATTTAATACATTGGAAAGGGGAAAAAAGGACCTTCA 804

Qy     489 nMetSerGlyGlyGlnLysGlnArg 497
           |     |||    |||    :::
Db     805 AGGAATGGAAGGGCCAAAACCCAAG 829
```

RESULT 8
BM904842
LOCUS   BM904842   998 bp   mRNA   linear   EST 12-MAR-2002
DEFINITION AGENCOURT_6699581 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5557655
            5', mRNA sequence.
ACCESSION   BM904842
VERSION   BM904842.1 GI:19355221
KEYWORDS   EST.
SOURCE   human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 998)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT   Contact: Robert Strausberg, Ph.D.

```
          Email: cgapbs-r@mail.nih.gov
          Tissue Procurement: ATCC/DCTD/DTP
           cDNA Library Preparation: Life Technologies, Inc.
           cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
           DNA Sequencing by: Agencourt Bioscience Corporation
           Clone distribution: MGC clone distribution information can be
          found through the I.M.A.G.E. Consortium/LLNL at:
          http://image.llnl.gov
          Plate: LLAM12279  row: f  column: 24
          High quality sequence stop: 738.
FEATURES             Location/Qualifiers
     source          1. .998
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5557655"
                     /clone_lib="NIH_MGC_72"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      271 a    223 c    248 g    256 t
ORIGIN

Alignment Scores:
Pred. No.:              3.24e-80      Length:        998
Score:                  834.00        Matches:       164
Percent Similarity:     74.12%        Conservative:  68
Best Local Similarity:  52.40%        Mismatches:    77
Query Match:            13.54%        Indels:        4
DB:                     14            Gaps:          0

US-09-873-409-5 (1-1222) x BM904842 (1-998)

Qy     745 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 764
           |||::::::|||||||||:::    :::|||:::||||||    |||||||||    ||||||
Db      61 GATGTGAGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCC 120

Qy     765 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 784
              |||     ||||||:::::||||||     |||||||||:::    |||:::|||||||||
Db     121 AATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAAT 180

Qy     785 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 804
                    |||:::|||        :::||||||||||||||||||||||||||::::::|||
Db     181 ATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTG 240

Qy     805 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 824
           |||:::|||:::|||    |||:::::::|||:::     |||::::::|||         :::
Db     241 TTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTG 300

Qy     825 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 844
           :::|||    |||    |||||||||:::||||||:::    ||||||||||||||||||
Db     301 TCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAA 360

Qy     845 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 864
           |||:::||||||    ||||||:::|||||||||||||:::|||:::      ||||||    |||
Db     361 GCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATG 420

Qy     865 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 884
           |||    :::    ||||||         :::|||||||:::    :::|||||||    |||    |||
Db     421 TATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGA 480

Qy     885 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 904
              ::::::|||:::    |||    :::|||||||:::||||||       |||||||||
Db     481 ATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTT 540

Qy     905 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 924
           ||||||||||||:::                |||:::    |||    :::    :::||||||:::
Db     541 GGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCA 600

Qy     925 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 944
           |||:::    :::||||||||||||:::|||:::               ||||||:::|||:::
Db     601 GCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCC 660

Qy     945 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 964
           |||||||||       |||||||||:::       ::::::||||||    |||    ||||||
Db     661 AAAGCCAAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGAC 720

Qy     965 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 984
           |||    |||    ||||||          |||:::|||    |||||||||:::    |||
Db     721 AGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGGTGA 780

Qy     985 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlySer-LeuSe 1004
           |||    |||    ||||||      ||||||||||:::    :::|||:::||||||||| |||
Db     781 GTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCCTGGA 840

Qy    1004 rIleGluArgGlyLysThrValAlaPheVal-GlySerSerGlyCysGlyLysSerThrS 1024
             ::::::::|||:::|||:::|||    ||| |||||||||||||||||||||||||
Db     841 GGTGAAGAAGGGGCAGACCCTGGCTCTGGTGGGGCAGCAGTGGCTGTGGGAAGAGCACAG 900

Qy    1024 er-ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 1043
                ||||||||||||:::|||:::|||:::|||:::        ::::::|||    ||||||
Db     901 TGGGTCCAGCTCCTGGAGCGGGTCTACAACCCCTTGGCAAGGAAAATGCTGCTTGATGGC 960

Qy    1044 ValAspAlaLys-GluLeuAsnValGlnTrp 1053
                     |||    |||:::|||    |||
Db     961 AAAAGAATAAAACCGACTGGATGTTCCATGG 991

RESULT 9
AK014319
LOCUS       AK014319             2676 bp    mRNA    linear   HTC 19-JAN-2002
DEFINITION  Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length
            enriched library, clone:3222401P09:ATP-binding cassette, sub-family
            B (MDR/TAP), member 8, full insert sequence.
ACCESSION   AK014319
VERSION     AK014319.1  GI:12852089
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (strain:C57BL/6J) 14, 17 days embryo head cDNA to
            mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:3222401P09.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
```

```
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409 (6821), 685-690 (2001)
  MEDLINE   21085660
   PUBMED   11217851
REFERENCE   5  (bases 1 to 2676)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C.,
            Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,
            Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F.,
            Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,
            Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,
            Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
            Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,
            Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and
            Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Gemomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. Second
            strand cDNA was prepared with the primer adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cleaved
            with BamHI and XhoI. cDNA of size comprised between 3 and 7 kb was
            selected before cloning. Vector: a modified pBluescript KS(+) after
            bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'
            end: BamHI. Host: DH10B.
FEATURES             Location/Qualifiers
     source          1. .2676
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:3222401P09"
                     /db_xref="MGD:MGI:1907488"
                     /db_xref="taxon:10090"
                     /clone="3222401P09"
                     /tissue_type="head"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="14, 17 days embryo"
     gene            1. .2676
                     /gene="Abcb8"
     CDS             72. .2225
                     /gene="Abcb8"
                     /note="ATP-binding cassette, sub-family B (MDR/TAP),
                     member 8
                     data source:MGD, source key:MGI:1351667, evidence:ISS
                     putative"
                     /codon_start=1
                     /protein_id="BAB29270.1"
                     /db_xref="GI:12852090"
                     /db_xref="MGD:MGI:1351667"
                     /translation="MLVHLFRFGIRGGPVPGWSLQSLRFQTFSAARSSDDRLSSHLLR
                     TVAQLRVQLRAHLPRAPPASHWSPSAWCWVGGTLVVPAVLWQHPRLCLIALCEAKESP
                     PAQPTRAPELRFNWKLFWHFLHPHLLALGAAIVLALGAALVNVQIPLLLGQLVEIVAK
                     YTRDHMGSFVSESRKLSVQLLLLYGVQGLLTFGYLVLLSHIGERMAMDMRKALFSSLL
                     RQDIAFFDAKKTGQLVSRLTTDVQEFKSSFKLVISQGLRSCTQVIGSLVSLSMLSPRL
                     TLMLAVVTPALMGVGTLMGSGLRKLSRQCQEQIARATGVADEALGNVRTVRAFAMEKR
                     EEERYQAELESCCCKAEELGRGIALFQGLSNIAFNCMVLGTLFIGGSLVAGQQLKGGD
                     LMSFLVASQTVQRSMASLSVLFGQVVRGLSAGARVFEYMALSPVIPLTGGYCIPNKDI
                     RGSITFQNVTFSYPCRPGFNVLKDFTLKLPSGKIVALVGQSGGGKTTVASLLERFYDP
                     EAGSVTLDGHDLRTLNPSWLRGQVIGFISQEPVLFATTIMENIRFGKLDASDEEVYTA
                     AREANAHEFISSFPDGYSTVVGERGTTLSGGQKQRLAIARALIKQPTVLILDEATSAL
                     DAESERVVQEALDRASAGRTVLVIAHRLSTVRAAHSIIVMANGQVCEAGTHEELLKKG
                     GLYSELIRRQTLDASLTSTPPAEKPEDPKSCQSKA"
     polyA_signal    2653. .2658
                     /gene="Abcb8"
                     /note="putative"
     polyA_site      2676
                     /gene="Abcb8"
                     /note="putative"
BASE COUNT      544 a    785 c    756 g    591 t
ORIGIN

Alignment Scores:
Pred. No.:              2.16e-79      Length:          2676
Score:                  833.50        Matches:         222
Percent Similarity:     53.82%        Conservative:    109
Best Local Similarity:  36.10%        Mismatches:      205
Query Match:            13.53%        Indels:          79
DB:                     11            Gaps:            20

US-09-873-409-5 (1-1222) x AK014319 (1-2676)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            ::::::|||   :::   |||:::|||||||||                  :::||||||
Db      471 ATTGTGCTGGCCTTAGGTGCGGCACTAGTGAAT------------GTGCAGATCCCCTTG 518

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIleSer------------------GlyCysLeu 34
            :::||||||::::::   :::   :::                        |||
Db      519 CTCCTGGGCCAGCTGGTGGAGATTGTCGCCAAGTACACGAGGGACCACATGGGGAGTTTC 578

Qy       35 ValGlnThrAsnThrTyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyVal 54
            |||      :::               :::|||   |||   |||         ||||||
Db      579 GTGTCTGAGTCCCGTAAGCTCAGCGTCCAGCTGCTCCTACTGTAC---------GGTGTT 629

Qy       55 AlaAlaLeuIle---PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArg 73
                  |||:::   |||||||||:::   :::         |||            |||
Db      630 CAGGGACTGCTGACCTTTGGATACCTAGTGCTGCTGTCCCACATT------GGTGAGCGC 683

Qy       74 GlnThrLysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrp 93
                        :::||||||      |||   |||:::|||   |||||||||   :::
Db      684 ATGGCCATGGACATGCGGAAAGCCCTTTTCAGCTCCCTGCTCCGGCAAGACATTGCTTTC 743

Qy       94 PheAspSerCysAspIleGlyGluLeuAsnThrArgMet---ThrAspIleAspLys--- 111
            ||||||:::         |||:::|||   :::|||:::   ||||||:::   :::
Db      744 TTTGATGCCAAAAAGACAGGGCAGCTAGTGAGTCGCTTGACTACTGATGTGCAAGAATTC 803

Qy      112 ---------------------IleSerAspGlyIleGlyAspLysIleAlaLeuLeuPhe 124
                                 ||||||   |||:::
Db      804 AAGTCATCCTTCAAGCTTGTCATCTCCCAGGGACTG------------------------ 839

Qy      125 GlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu---ValLysGlyTrpLys 143
                        |||      ||||||      |||   |||   :::         :::
Db      840 ---CGCAGCTGCACCCAGGTGATTGGTAGCCTGGTGTCCCTGTCTATGCTGTCCCCTCGC 896

Qy      144 LeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSer 163
```

```
            ||||||||:::                   :::|||    :::|||
Db     897 CTTACCCTGATGCTGGCTGTCGTCACACCCGCCCTCATGGGAGTGGGCACCCTGATGGGC 956

Qy     164 ArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaVal 183
              :::         |||:::    :::                :::::|||      |||
Db     957 TCAGGCCTCCGAAAGCTGTCTCGCCAGTGTCAGGAGCAGATTGCCAGGGCAACAGGTGTA 1016

Qy     184 AlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGlu 203
           |||:::|||   |||    :::::||||||||    ||||||       :::::::::|||
Db    1017 GCAGATGAGGCCCTTGGCAATGTTCGGACTGTGCGGGCCTTCGCGATGGAGAAGAGGGAG 1076

Qy     204 LeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrp 223
              :::|||                                             :::
Db    1077 GAGGAACGC---------------------------------------TATCAAGCAGAA 1097

Qy     224 LeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAsp 243
           |||   ||||||                                   |||::::::
Db    1098 CTGGAGTCATGCTGCTGT------------------------------AAAGCAGAAGAA 1127

Qy     244 PheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhe------ 261
           |||         |||   ||||||             |||       :::   |||
Db    1128 CTGGGC------AGGGGCATCGCCTTGTTCCAAGGGCTCTCCAACATCGCTTTCAACTGT 1181

Qy     262 PheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGly 281
              :::   ||||||                 :::   |||   ||||||:::
Db    1182 ATGGTCTTGGGCACC------------CTGTTCATTGGGGGCTCCCTTGTGGCTGGACAG 1229

Qy     282 Glu---ProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
           :::      |||      :::   :::   |||             :::|||      |||
Db    1230 CAGCTGAAAGGGGGAGACCTCATGTCCTTCCTGGTGGCTTCCCAGACAGTACAGAGGTCT 1289

Qy     301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGly--- 319
                       :::                                :::   ||||||
Db    1290 ATGGCCAGCCTCTCTGTCCTGTTTGGTCAG---------------GTGGTACGTGGGCTG 1334

Qy     320 ---AlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
              |||         :::|||:::   :::                |||   |||
Db    1335 AGTGCGGGAGCCCGAGTCTTCGAATACATGGCCCTGAGCCCTGTCATCCCA------TTG 1388

Qy     339 ThrAlaGlyTyr------LysProGluSerIleGluGlyThrValGluPheLysAsnVal 356
           |||   ||||||                :::      |||   |||::::::   |||:::||||||
Db    1389 ACCGGGGGCTACTGCATCCCCAACAAGGACATTCGTGGTTCCATCACCTTCCAAAATGTC 1448

Qy     357 SerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIle 376
           :::|||:::||||||      ||||||           :::||||||          |||::::::
Db    1449 ACCTTCAGCTACCCCTGCAGACCTGGCTTCAATGTGCTCAAGGACTTCACCCTGAAGCTG 1508

Qy     377 LysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValVal 396
              ||||||:::   |||||||||||||||      :::|||   ||||||:::||||||
Db    1509 CCCTCTGGCAAGATTGTGGCTCTTGTGGGCCAGTCTGGGGGAGGAAAGACCACAGTTGCC 1568

Qy     397 GlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAsp 416
              ||||||:::|||   |||||||||:::      |||    :::   :::|||    :::|||
Db    1569 TCCTTGCTGGAACGCTTCTATGACCCTGAAGCTGGCTCGGTGACGTTGGATGGGCATGAC 1628

Qy     417 IleArgAlaLeuAsnValArgHisTyrArgAspHis---IleGlyValValSerGlnGlu 435
           :::|||   ||||||                |||         ||||||   :::|||||||||
Db    1629 CTGCGAACTCTCAACCCCTCCTGGCTCCGGGGCCAGGTCATAGGTTTCATCAGCCAGGAG 1688

Qy     436 ProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThr 455
           ||||||||||||   |||||||||         ||||||::::::|||:::   |||    :::
Db    1689 CCAGTCCTGTTTGCAACAACCATCATGGAGAATATCCGATTTGGGAAGCTGGATGCTTCC 1748

Qy     456 AspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhe 475
           |||||||||:::         |||||||||||||||||||||::::::||||||       |||
Db    1749 GATGAAGAGGTGTACACAGCTGCACGAGAAGCCAATGCCCACGAGTTCATCAGCAGCTTC 1808

Qy     476 ProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLys 495
           |||:::   ::::::|||:::|||||||||:::|||      :::||||||||||||||||||
Db    1809 CCCGATGGCTACAGCACTGTGGTTGGTGAGCGGGGCACAACCTTGTCTGGTGGCCAGAAG 1868

Qy     496 GlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGlu 515
           ||||||:::||||||||||||||||||::::::   |||   :::||||||||||||||||||

Db    1869 CAGCGCCTAGCCATCGCACGTGCCCTCATCAAGCAGCCCACAGTGCTGATCCTGGACGAG 1928

Qy     516 AlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAla 535
           ||||||||||||||||||:::||||||:::         ||||||   ||||||::::::|||
Db    1929 GCCACCAGTGCGCTAGATGCAGAATCCGAGAGGGTGGTACAGGAGGCCCTGGACCGGGCC 1988

Qy     536 SerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAsp 555
           |||      ||||||||    :::|||:::||||||||||||||||||:::|||:::|||
Db    1989 AGTGCTGGCCGCACCGTGTTGGTCATTGCCCACCGGCTTAGTACTGTCCGTGCAGCCCAC 2048

Qy     556 LeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMet 575
              |||:::   :::   :::|||   :::   |||   |||   |||   ||||||:::
Db    2049 TCCATCATTGTCATGGCCAATGGCCAAGTCTGTGAGGCTGGGACCCACGAAGAACTCCTT 2108

Qy     576 AlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle 590
              |||   ||||||||         |||:::         |||   :::
Db    2109 AAAAAAGGCGGGCTCTATTCAGAGCTTATCCGGAGACAAACCCTG 2153

RESULT 10
BF584668
LOCUS       BF584668                 894 bp    mRNA    linear   EST 12-DEC-2000
DEFINITION  602098406F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4218385 5',
            mRNA sequence.
ACCESSION   BF584668
VERSION     BF584668.1  GI:11658386
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 894)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9798  row: 1  column: 02
            High quality sequence stop: 651.
FEATURES             Location/Qualifiers
     source          1..894
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4218385"
                     /clone_lib="NCI_CGAP_Co24"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.6 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      263 a    203 c    242 g    185 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:               2.18e-73       Length:          894
Score:                   771.00         Matches:         171
Percent Similarity:      79.72%         Conservative:    57
Best Local Similarity:   59.79%         Mismatches:      48
Query Match:             12.52%         Indels:          15
DB:                      12             Gaps:            3

US-09-873-409-5 (1-1222) x BF584668 (1-894)

Qy     293 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 312
           |||||||||||||||:::         ::::::   :::|||   |||   |||:::   |||
Db       5 GTCTTCTTTTCCGTGTTAAT-GGAGCATTCAGTGTTGGACAGGCATCTCCAAATATTGAA 63
```

```
Qy     313 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 332
              |||||   |||||||||||||:::   :::|||:::::|||||   |||||
Db      64 GCCTTCGCCAATGCACGAGGAGCAGCTTATGAAGTCTTCAAAATAATTGATAATAAGCCC 123

Qy     333 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 352
           |||||||:::|||||   :::|||:::|||||::::::|||:::|||   :::|||
Db     124 AGTATAGACAGCTTCTCAAAGAGTGGGCACAAACCAGACAACATACAAGGAAATCTGGAA 183

Qy     353 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 372
           ||||||||:::   |||:::||||||||||      ::::::||||||||||||||
Db     184 TTTAAGAATATTCACTTCAGTTACCCATCTCGAAAAGAAGTTCAGATCTTGAAGGGCCTC 243

Qy     373 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 392
           ||||||::::::||||||||:::|||||||||||||||||   :::|||   ||||||
Db     244 AATCTGAAGGTGAAGAGCGGACAGACGGTGGCCCTGGTTGGCAACAGTGGCTGTGGAAAA 303

Qy     393 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 412
              ||||||   |||||||||:::||||||||||||||||||   ||||||   :::   :::
Db     304 AGCACAACTGTCCAGCTGATGCAAAGGCTCTACGACCCCCTAGATGGCATGGTCAGTATC 363

Qy     413 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 432
           |||   ||||||||, :::|||||||||:::   |||:::   |||||||||||
Db     364 GACGGACAGGACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCATTGGTGTGGTG 423

Qy     433 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 452
           ||||||||||||||||||||   ||||||||||:::   ||||||:::|||||||||:::
Db     424 AGTCAGGAACCTGTGCTGTTTGCCACCACGATCGCCGAGAACATTCGCTATGGCCGAGAA 483

Qy     453 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 472
           |||||||||   :::|||:::|||:::|||   :::|||||||||||||||||||||||||
Db     484 GATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTCATC 543

Qy     473 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 492
           |||:::   |||::::::|||:::|||||||||||||||:::|||||||||:::||||||
Db     544 ATGAAACTGCCCCACCAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGG 603

Qy     493 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 512
           ||||||||||||||||||   |||||||||||||||||||||||||||||   |||
Db     604 GGACAGAAACAGAGAATCGC-ATTGCCCGGGCCCTGGTCCGCAATCCCAAGTGCCTTTC- 661

Qy     513 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla-Le 532
           :::|||||||||||||||   ||||||:::||||||::::::   ||||||||||| ||
Db     662 GTGGACGAAGCCACCTCAGC-CTGGATACAGAAAGTGAAGCTGGGGTTCAGGCCGCACCT 720

Qy     532 uGluLysAlaSerLysGlyArgThrThrIle------------ValValAlaHisArgLe 548
           |::::::         ||||||   ||||||:::          ::::::       :::||
Db     721 AGATAGGGTACAAAAGGGCNGGACCACATTGGGAAGCTCATCGCTGTTACCGTCGAAGCT 780

Qy     548 uSerThrIleArgSerAlaAspLeuIleVal------ThrLeuLysAspGlyMetLeuAl 566
           |:::   :::|||                :::|||   |||||||||          ::
Db     781 GACGGCCATGCGGCCGAAGGGGGG-CTTGTGGGAACGACTCTGAAAAGC---------TC 830

Qy     566 aGluLysGlyAlaHis 571
           :|||||||||   |||
Db     831 TGAAAAAGGTACTCAT 846
```

```
RESULT 11
BG248052
LOCUS       BG248052               1019 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  602359987F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4488404 5',
            mRNA sequence.
ACCESSION   BG248052
VERSION     BG248052.1  GI:12757867
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1019)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10334  row: n  column: 21
            High quality sequence stop: 650.
FEATURES             Location/Qualifiers
     source          1. .1019
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4488404"
                     /clone_lib="NCI_CGAP_Mam1"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="10 months, virgin"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
BASE COUNT      276 a    262 c    283 g    196 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              3.13e-73      Length:        1019
Score:                  770.50        Matches:       154
Percent Similarity:     79.17%        Conservative:  36
Best Local Similarity:  64.17%        Mismatches:    47
Query Match:            12.51%        Indels:        3
DB:                     12            Gaps:          1

US-09-873-409-5 (1-1222) x BG248052 (1-1019)

Qy     977 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 996
           ||||||||::::::|||      |||   |||   ||||||   ||||||      :::
Db       4 GAAGGAAATGTGCAATTTAGTGGAGTCGTGTTCAACTATCCCACCCGACCCAGCATCCCA 63

Qy     997 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 1016
           :::|||:::||||||||||||   :::::::::|||:::|||:::|||   |||||||||
Db      64 GTGCTTCAGGGGCTGAGCCTTGAGGTGAAGAAGGGCCAGACGCTGGCCCTGGTGGGCAGC 123

Qy    1017 SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGln 1036
           ||||||||||||||||||||   ||||||||||||:::|||||||||||||||||:::
Db     124 AGTGGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAGCGCTTCTACGACCCCATGGCT 183

Qy    1037 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 1056
           |||   |||   ||||||   :::   |||:::||||||||||||||||||||||:::
Db     184 GGATCAGTGTTTCTAGATGGCAAAGAAATAAAGCAACTGAATGTCCAGTGGCTCCGAGCA 243

Qy    1057 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 1076
           |||:::   ||||||   |||||||||:::||||||:::|||||||||||||||||||||
Db     244 CAGCTGGGCATTGTGTCCCAAGAGCCCATTCTCTTTGACTGCAGCATCGCAGAGAACATT 303

Qy    1077 AlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla 1096
           |||||||||||||||||||||||||||         :::||||||      ||||||
Db     304 GCCTACGGAGACAACAGCCGGGTCGTGTCTTATGAGGAGATTGTGAGGGCAGCCAAGGAG 363

Qy    1097 AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeu 1116
           ||||||||||||   ||||||:::   ||||||:::||||||||||||:::||||||
Db     364 GCCAACATCCACCAGTTCATCGACTCGCTACCTGATAAATACAACACCAGAGTAGGAGAC 423

Qy    1117 LysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeu 1136
           ||||||   |||||||||||||||||||||||||||:::||||||||||||||||||:::
Db     424 AAAGGCACTCAGCTGTCGGGTGGGCAGAAGCAGCGCATCGCCATCGCACGCGCCCTCGTC 483

Qy    1137 GlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGlu 1156
           ::::::|||   ||||||||||||||||||||||||||||||||||||   :::||||||
Db     484 AGACAGCCTCACATTTTACTTCTGGACGAAGCAACATCAGCTCTGGATACAGAAAGTGAA 543
```

```
Qy   1157 LysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThr 1176
          |||||||||||  ||||||||||||||||  ||||||||||:::|||:::
Db    544 AAGGTTGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACCTGCATTGTGATCGCT 603

Qy   1177 HisArgLeuSerAlaIleGlnAsnAlaAspLeuIle-ValValLeuHis-AsnGlyLysI 1196
          |||||               |||||||||||||||| ||||||:::    |||   |||:
Db    604 CACCGCTGTCCACCATCCCAGAACGCGGACTTGATCGGTGGTGATTCAAGAACCGGAAGG 663

Qy   1196 leLys---GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1213
          ::|||      :::          |||:::::::::     :::   ||||||||
Db    664 TCAAGGGAGCACGGGCACCCCAACAGCAAGTGGTGGCGCAGAAGGGCATCTACTTC 719

RESULT 12
BG293345
LOCUS       BG293345                  726 bp    mRNA    linear   EST 21-FEB-2001
DEFINITION  602390738F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4502552 5',
            mRNA sequence.
ACCESSION   BG293345
VERSION     BG293345.1  GI:13052943
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 726)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10371  row: 1  column: 09
            High quality sequence stop: 658.
FEATURES             Location/Qualifiers
     source          1. .726
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4502552"
                     /clone_lib="NIH_MGC_94"
                     /tissue_type="retina"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 3.3 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      189 a    220 c    190 g    127 t
ORIGIN

Alignment Scores:
Pred. No.:               1.42e-71      Length:        726
Score:                   753.00        Matches:       152
Percent Similarity:      84.29%        Conservative:  25
Best Local Similarity:   72.38%        Mismatches:    30
Query Match:             12.23%        Indels:        3
DB:                      12            Gaps:          0

US-09-873-409-5 (1-1222) x BG293345 (1-726)

Qy   1011 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 1030
          :::|||  |||||||||||||||||||||||||||||    |||||||||||:::|||
Db      2 CTGGCCCTGGTGGGCAGCAGTGGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAGCGC 61

Qy   1031 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 1050
             |||||||||:::   |||   ||||||   ||||||    :::||||||:::||||||
Db     62 TTCTATGACCCCATGGCTGGATCAGTGCTCTTAGATGGTCAAGAAGCAAAGAAACTCAAT 121

Qy   1051 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 1070
          ||||||||||||||:::|||:::   ||||||   |||||||||:::||||||:::|||
Db    122 GTCCAGTGGCTCCGAGCTCAACTGGGCATTGTGTCCCAGGAACCCATTCTCTTTGACTGC 181

Qy   1071 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 1090
          |||||||||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db    182 AGCATCGCAGAGAACATCGCCTATGGAGACAACAGCCGGGTCGTGCCTCATGATGAGATT 241

Qy   1091 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 1110
                ||||||      ||||||||||||   |||||||||   ||||||:::||||||
Db    242 GTGAGGGCAGCCAAGGAGGCCAACATCCACCCCTTCATCGAGACGCTGCCCCAAAAATAT 301

Qy   1111 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 1130
          ||||||:::||||||   ||||||   ||||||||||||||||||||||||||||:::|||
Db    302 AACACAAGAGTAGGAGACAAGGGGACGCAGCTCTCTGGGGGCCAGAAGCAGAGGATTGCC 361

Qy   1131 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 1150
          ||||||||||||:::::::::|||::::::||||||||||||||||||||||||||||||
Db    362 ATCGCCCGAGCCCTCATCAGACAGCCTCGGGTCCTACTGCTGGATGAAGCCACGTCAGCT 421

Qy   1151 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 1170
          ||||||   :::|||||||||||||||||||   ||||||||||||||||||   ||||||
Db    422 CTGGATACTGAGAGTGAAAAGGTTGTCCAGGAAGCACTGGACAAAGCCAGGGAAGGCCGC 481

Qy   1171 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla-AspLeuIleValVa 1190
          ||||||:::   :::   ||||||||||||   ||||||||||||   ||||||||||||
Db    482 ACCTGCATTGCGATCGCTCACCGCCTGTCCACCATCCAGAACGCGGCACTTGATCGTGGT 541

Qy   1190 lLeuHisAsnGlyLysIleLysGluGlnGlyThr-HisGlnGluLeuLeuArgAsnArgA 1210
          |:::   ||||||||||:::||||||    ||||||  ||||||:::||||||      :::
Db    542 GATTGAGAACGGCAAGGTCAAGGAGCACGGCACCCCACCAGCAGCTGCTGGCGCAGAAGG 601

Qy   1210 spIleTyrPhe-LysLeuValAsn 1217
            ||||||||| :::|||||||||
Db    602 GCATCTATTTCTCAACTGGTCAAC 625

RESULT 13
AZ682350
LOCUS       AZ682350                  871 bp    DNA     linear   GSS 14-DEC-2000
DEFINITION  ENTKB16TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   AZ682350
VERSION     AZ682350.1  GI:11819496
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
  ORGANISM  Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 871)
  AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Forward
            Class: shotgun
            High quality sequence start: 16
            High quality sequence stop: 860.
FEATURES             Location/Qualifiers
     source          1. .871
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
```

```
            Institute for Genomic Research (TIGR), Rockville, MD.
            Genomic DNA isolated from broth cultures of E. histolytica
            using a  method described by Clark and Diamond (Clark,
            C.G., and Diamond, L.S.  (1993) Entamoeba histolytica: a
            method for isolate identification.  Exp. Parasitol.
            77:450.). The DNA was mechanically sheared to give a
            tight size distribution (~2 kb). The v + i method used for
            the library  construction is described in detail in Smith,
            H.O. and Venter, J.C.  (Making small insert libraries for
            whole genome shotgun sequencing  projects. In Genome
            Sequencing: A Practical Approach, eds. M. Vaudin  and B.
            Barell, Oxford University Press, 1999)."
BASE COUNT      349 a    117 c    174 g    231 t
ORIGIN

Alignment Scores:
Pred. No.:               2.23e-71      Length:         871
Score:                   752.50        Matches:        151
Percent Similarity:      71.64%        Conservative:   46
Best Local Similarity:   54.91%        Mismatches:     75
Query Match:             12.22%        Indels:         3
DB:                      17            Gaps:           1

US-09-873-409-5 (1-1222) x AZ682350 (1-871)

Qy    317 AlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsn 336
          |||:::   |||||||||:::::::::|||  ||||||:::   |||   |||||
Db     45 GCTAAAGTTGCCGCATTTAATGTTTATCAAACAATTGATAGAATCCCAGATATTGATTGT 104

Qy    337 PheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnVal 356
             |||      |||      |||                |||   :::   |||::::::|||
Db    105 CAGTCTATTGGAGGTGAATGTCCAACTGAGTGTAATGGAAATATTAGATTTGAGGATGTT 164

Qy    357 SerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIle 376
             |||   ||||||:::|||   |||        :::||||||||||||:::|||   |||
Db    165 CAATTTGTTTATCCAACAAGACTGTCTCATCATGTATTAAAAGGACTTGACCTTGAAATT 224

Qy    377 LysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValVal 396
          |||   |||:::|||:::|||||||||||     :::|||   |||||||||||     :::
Db    225 AAGAAAGGACAAACAATTGCATTAGTTGGAGCATCAGGATGTGGGAAGTCAACTACTATT 284

Qy    397 GlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAsp 416
          ||||||:::||||||   |||||||||:::   |||   :::   :::|||      |||
Db    285 CAATTAATCCAAAGAAATTATGATCCAAATGGTGGAAGAGTAACATTAGACGGAAAAGAT 344

Qy    417 IleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluPro 436
          ||||||   ||||||::::::         |||:::   ||||||:::|||   |||||||||
Db    345 ATACGAGAGTTGAATATCAAATGGTTAAGAAATCAAATAGGATTAGTTGGACAAGAACCA 404

Qy    437 ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAsp--------- 453
          |||||||||      ||||||      ||||||               :::
Db    405 GTGTTGTTTGCAGGAACAATTCGAGAAAATATTATGCTTGGAGCTAAAGAAGGAGCAACA 464

Qy    454 ValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMet 473
             ::::::|||||||||         |||:::   |||||||||:::|||||||||
Db    465 CCAAGTGAAGAAGAGATGATTGAATGTGCTAAAATGGCAAATGCACATGACTTCATTTCT 524

Qy    474 GluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGly 493
          :::   |||      ::::::|||::::::|||||||||||||||      :::|||||||||
Db    525 AAACTTCCAGAAGGATATGACACAATAATTGGAGAAAAAGGAGCATTATTATCAGGAGGA 584

Qy    494 GlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeu 513
          ||||||||||||||||||||||||||||||:::|||||||||   ||||||:::|||
Db    585 CAAAAACAAAGAATTGCAATTGCACGTGCATTGATTCGAAACCCATCTATTCTTCTTCTT 644

Qy    514 AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGlu 533
          |||||||||||||||||||||||||||::::::|||:::         ||||||   |||||||||
Db    645 GATGAAGCTACATCAGCACTTGATACACAAAGTGAAAAGATTGTACAAGAAGCACTTGAA 704

Qy    534 LysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSer 553
          ||||||||||||||||||||||||:::|||||||||||||||:::|||:::|||:::
Db    705 AAAGCATCTAAAGGAAGAACAACAATTATTGTAGCACATAGACTAACAACTGTTAGAAAT 764

Qy    554 AlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGlu 573
          ||||||   |||            |||   :::   |||:::|||   |||   |||
Db    765 GCAGATAAAATATGTGTATTTCATCAAGGAGAAATTATTGAACAAGGAAAACATCAAGAA 824

Qy    574 LeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln 588
          ||||||      :::|||   ||||||   ||||||      |||
Db    825 TTAATGGATTTGAAAGGAACATATTATGGATTAGTCAAAAGACAA 869

RESULT 14
BF796582
LOCUS       BF796582               944 bp    mRNA     linear   EST 12-JAN-2001
DEFINITION  602258463F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4341710 5',
            mRNA sequence.
ACCESSION   BF796582
VERSION     BF796582.1  GI:12101636
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 944)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis Staudt, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9955  row: n  column: 15
            High quality sequence start: 9
            High quality sequence stop: 669.
FEATURES             Location/Qualifiers
     source          1. .944
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4341710"
                     /clone_lib="NIH_MGC_85"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.867 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      265 a    241 c    237 g    201 t
ORIGIN

Alignment Scores:
Pred. No.:               3.76e-71      Length:         944
Score:                   751.00        Matches:        166
Percent Similarity:      74.62%        Conservative:   31
Best Local Similarity:   62.88%        Mismatches:     60
Query Match:             12.19%        Indels:         9
DB:                      12            Gaps:           2

US-09-873-409-5 (1-1222) x BF796582 (1-944)

Qy    943 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 962
          |||:::|||||||||         |||||||||||||||   |||   |||::::::|||
Db     39 TATGCTAAAGCTAAGCTGTCTGCAGCCCACTTATTCATGCTGTTTGAAAGACAACCTCTG 98

Qy    963 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 982
          |||||||||   |||:::||||||   |||||||||         |||||||||:::   |||
Db     99 ATTGACAGCTACAGTGAAGAGGGGCTGAAGCCTGATAAATTTGAAGGAAATATAACATTT 158

Qy    983 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002
             ||||||   |||   ||||||   |||   :::|||   :::|||:::|||||||||
Db    159 AATGAAGTCGTGTTCAACTATCCCACCCGAGCAAACGTGCCAGTGCTTCAGGGGCTGAGC 218
```

```
Qy   1003 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 1022
          |||   :::::::::|||:::|||:::|||   |||||||||||||||||||||||||
Db    219 CTGGAGGTGAAGAAAGGCCAGACACTAGCCCTGGTGGGCAGCAGTGGCTGTGGGAAGAGC 278

Qy   1023 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1042
          |||   |||||||||||:::|||   ||||||||:::   |||   ||||||   |||
Db    279 ACGGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCGGGGACAGTGCTTCTCGAT 338

Qy   1043 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 1062
          |||   :::||||||:::||||||||||||||||||||||:::|||:::   ||||||
Db    339 GGTCAAGAAGCAAAGAAACTCAATGTCCAGTGGCTCAGAGCTCAACTCGGAATCGTGTCT 398

Qy   1063 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082
          |||||||||:::||||||:::||||||||||||||||||||||||||||||||||||||||
Db    399 CAGGAGCCTATCCTATTTGACTGCAGCATTGCCGAGAATATTGCCTATGGAGACAACAGC 458

Qy   1083 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1102
          |||||||||   |||||||||   ||||||   ||||||||||||||||   |||
Db    459 CGGGTTGTATCACAGGATGAAATCGTGAGTGCAGCCAAAGCTGCCAACATACATCCTTTC 518

Qy   1103 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122
          ||||||   ||||||   ||||||   |||:::||||||   ||||||   |||||||||
Db    519 ATCGAGACGTTACCCCACAAATATGAAACAAGAGTGGGAGATAAGGGGACTCAGCTCTCA 578

Qy   1123 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLys-ProLysIleLe 1142
          ||||||   |||||||||:::|||||||||||||||||||:::::::::  |||:::|||||
Db    579 GGAGGTCACAAACAGAGGATTGCTATTGCCCGAGCCCTCATCAGACAACCCTCAAATCCT 638

Qy   1142 uLeuLeuAspGluAlaThrSerAla-LeuAspAsnAspSerGluLysValVal------G 1160
          |||||||||||||||||||||||||  ||||||                |||        :
Db    639 CCTGTTGGATGAAGCTACATCAGCCTCTGGATACCTGACAAGTGAACACGGTTGTCCCAA 698

Qy   1160 lnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuS 1180
          ::                |||         |||||||||||||||:::   ||||||
Db    699 GAAGCCCCTGTCACACAGCAGAAGAGCGCCGCACCTGCCTTGTGATTGCTCACCGCTGTC 758

Qy   1180 erAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnG 1200
                ||||||||||||||   |||||||||         |||::::::|||
Db    759 AAC--ATCCAGAATGCAGACTT-ATAGTGGTGTTCCCAACGGGAGAAGTCAAG------- 808

Qy   1200 lyThrHis 1202
            ||||||
Db    809 --ACACAT 814
```

RESULT 15
AK020318

```
LOCUS       AK020318                872 bp    mRNA    linear   HTC 19-JAN-2002
DEFINITION  Mus musculus adult male epididymis cDNA, RIKEN full-length enriched
            library, clone:9230106F14:similar to MULTIDRUG RESISTANCE PROTEIN 3
            (P-GLYCOPROTEIN 3) (MDR1A), full insert sequence.
ACCESSION   AK020318
VERSION     AK020318.1  GI:12860872
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (strain:C57BL/6J) adult male epididymis cDNA to mRNA,
            clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:9230106F14.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409 (6821), 685-690 (2001)
  MEDLINE   21085660
   PUBMED   11217851
REFERENCE   5  (bases 1 to 872)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C.,
            Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,
            Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F.,
            Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,
            Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,
            Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
            Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,
            Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and
            Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Gemomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. cDNA went
            through one round of normalization to Rot = 10.0 and subtraction to
            Rot = 185.2. Second strand cDNA was prepared with the primer
```

```
            adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cleaved
            with BamHI and XhoI. Vector: a modified pBluescript KS(+) after
            bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'
            end: BamHI. Host: DH10B.
FEATURES             Location/Qualifiers
     source          1..872
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:9230106F14"
                     /db_xref="MGD:MGI:1911565"
                     /db_xref="taxon:10090"
                     /clone="9230106F14"
                     /sex="male"
                     /tissue_type="epididymis"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     CDS             86..>872
                     /note="data source:SPTR, source key:P21447, evidence:ISS
                     putative
                     similar to MULTIDRUG RESISTANCE PROTEIN 3 (P-GLYCOPROTEIN
                     3) (MDR1A)"
                     /codon_start=1
                     /protein_id="BAB32070.1"
                     /db_xref="GI:12860873"
                     /db_xref="MGD:MGI:1924956"
                     /translation="MANSERTNGLQETNQRYGPLQEQVPKVGNQAVGPIEIFRFADNL
                     DIVLMTLGILASMINGATVPLMSLVLGEISDHLINGCLVQTNRTKYQNCSQTQEKLNE
                     DIIVLTLYYIGIGAAALIFGYVQISFWVITAARQTTRIRKQFFHSILAQDISWFDGSD
                     ICELNTRMTGDINKLCDGIGDKIPLMFQNISGFSIGLVISLIKSWKLSLVVLSTSPLI
                     MASSALCSRMIISLTSKELDAYSKAGAVAEEALSSIQTVTAFGA"
BASE COUNT      262 a    178 c    187 g    245 t
ORIGIN

Alignment Scores:
Pred. No.:              1.69e-70        Length:          872
Score:                  744.50          Matches:         151
Percent Similarity:     80.37%          Conservative:    21
Best Local Similarity:  70.56%          Mismatches:      27
Query Match:            12.09%          Indels:          15
DB:                     11              Gaps:            2

US-09-873-409-5 (1-1222) x AK020318 (1-872)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          |||   ||||||||||||||||||:::::::||||||||     :::||||||||    |||
Db    230 ATGACTCTGGGAATATTAGCATCCATGATAAATGGAGCCACCGTTCCTTTAATGTCCCTG 289

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
          |||||||||||:::||||||:::||||||:::||||||||||||||||||||
Db    290 GTTTTAGGAGAAATAAGTGATCATTTAATTAATGGATGCCTAGTACAAACTAACAGAACT 349

Qy     41 SerPhe------------------------------------------PheArgLeuThr 46
                                                            ||||||
             :::
Db    350 AAATATCAGAACTGTTCTCAGACTCAAGAAAAGCTGAATGAAGATATCATTGTGTTGACT 409

Qy     47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
          |||||||||:::||||||||    ||||||||||||||||||||:::|||||||||
Db    410 CTATATTATATTGGAATAGGAGCAGCTGCCCTCATTTTTGGCTATGTACAGATTTCCTTC 469

Qy     67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
          |||:::||||||||||||||||||||    ||||||||||||||||||||||||:::
Db    470 TGGGTCATAACTGCAGCCCGGCAAACCACAAGAATCCGAAAACAGTTTTTTCATTCAATT 529

Qy     87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
          ||||||||||||||    ||||||||         ||||||    ||||||||||||||||||
Db    530 TTGGCACAAGACATCAGCTGGTTTGATGGCAGTGACATCTGTGAACTTAACACCCGCATG 589

Qy    107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
          |||   ||||||:::|||:::    ||||||||||||||||||||    |||:::||||||
Db    590 ACTGGTGACATCAACAAACTCTGTGATGGTATTGGAGATAAGATCCCTCTGATGTTTCAG 649

Qy    126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
          |||:::|||   |||||||||||||||   :::   |||:::|||   |||||||||:::
Db    650 AACATATCTGGGTTTTCTATTGGCCTGGTGATAAGTTTGATAAAAAGCTGGAAACTCTCC 709

Qy    146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
          ||||||    |||||||||||||||||||||||||||||:::|||    ||||||||||||
Db    710 CTGGTGGTTCTGTCTACATCTCCTCTCATAATGGCTTCATCGGCACTGTGCTCTAGGATG 769

Qy    166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
          :::||||||||||||||||||||||||    ||||||||||||||||||||||||||||||
Db    770 ATTATCTCATTGACCAGCAAGGAGCTGGATGCCTATTCCAAAGCTGGGGCTGTGGCTGAA 829

Qy    186 GluValLeuSerSerIleArgThrValIleAlaPheArgAla 199
          |||   ||||||||||||:::||||||    ||||||    |||
Db    830 GAAGCCTTGTCATCAATCCAAACAGTCACAGCCTTTGGAGCC 871

Search completed: March 31, 2003, 13:50:58
Job time : 8432.26 secs
```

```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 30, 2003, 01:20:01 ; Search time 1194.3 Seconds
                                           (without alignments)
                                           2304.231 Million cell updates/sec

Title:          US-09-873-409-5
Perfect score:  6159
Sequence:       1 MILGILASLVNGACLPLMPL..........QELLRNRDIYFKLVNAQSVQ 1222

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters:      4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115416_19216/app_query.fasta_1.7544
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_4247_@runat_27032003_115416_19216 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      N_Geneseq_101002:*
                1:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
                2:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
                3:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
                4:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
                5:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
                6:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
                7:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
                8:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
                9:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
                10:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
                11:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
                12:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
                13:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
                14:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
                15:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
                16:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
                17:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
                18:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
                19:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
                20:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
                21:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
                22:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
                23:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
                24:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5737	93.1	3699	24	ABK83223	Human transporter
2	3533.5	57.4	4279	22	AAD03506	Dog P-glycoprotein
3	3533	57.4	4195	22	AAF86128	Cynomologous monke
4	3530.5	57.3	3860	21	AAZ49332	Human wild-type mu
5	3530.5	57.3	3860	24	ABA94365	Human BCRP DNA rel
6	3530.5	57.3	4279	22	AAD03504	Dog P-glycoprotein
7	3530.5	57.3	4279	22	AAD03505	Dog P-glycoprotein
8	3525.5	57.2	4279	22	AAD03488	Dog P-glycoprotein
9	3525.5	57.2	4317	22	AAD03489	Dog P-glycoprotein
10	3524.5	57.2	4186	22	AAF86127	Cynomologous monke
11	3521.5	57.2	3860	21	AAZ49333	Human G185V mutant
12	3521.5	57.2	3860	24	ABA94366	Human BCRP DNA rel
13	3521.5	57.2	4349	22	AAH57442	Human intestine ce
14	3521.5	57.2	4646	21	AAZ94738	Human ATP binding
15	3521.5	57.2	4646	24	AAD38994	Human mdr1 gene.
16	3521.5	57.2	4646	24	ABL68592	Kidney cancer rela
17	3521.5	57.2	4646	24	ABL68880	Kidney cancer rela
18	3521.5	57.2	6505	17	AAT13394	Hybrid vector pSF-
19	3521.5	57.2	8630	21	AAZ24041	Retroviral M4 mdr-
20	3521.5	57.2	8630	21	AAZ24042	Retroviral vector
21	3520	57.2	4264	19	AAV66533	Mutated human P-gl
22	3520	57.2	4264	19	AAV66534	Mutated human P-gl
23	3519.5	57.1	4669	8	AAN70752	Sequence of human
24	3517	57.1	4425	21	AAZ52048	Rat multidrug resi
25	3515.5	57.1	4378	11	AAQ04522	Multidrug Resistan
26	3515	57.1	4369	21	AAZ52047	Rat multidrug resi
27	3514	57.1	3840	24	ABL91687	Human polynucleoti
28	3514	57.1	3988	21	AAZ88973	Human MDR-1 DNA.
29	3512.5	57.0	4669	14	AAQ52726	Sequence of human
30	3510.5	57.0	4646	15	AAQ72872	Human multidrug re
31	3503.5	56.9	4669	19	AAV32645	Human P glycoprote
32	3503.5	56.9	4669	24	ABK52041	cDNA encoding huma
33	3503	56.9	4189	21	AAZ49334	Murine multidrug r
34	3503	56.9	4189	24	ABA94367	Mouse BCRP DNA rel
35	3503	56.9	4313	14	AAQ38950	Mouse multidrug re
36	3499	56.8	4788	21	AAZ49335	Murine multidrug r
37	3499	56.8	4788	24	ABA94368	Mouse BCRP DNA rel
38	3494	56.7	4233	21	AAZ90198	Rat mdr1b2 (multis
39	3494	56.7	4233	22	AAF27498	Rat mdr1b2 multidr
40	3481	56.5	3924	21	AAZ94742	Human ATP binding
41	3481	56.5	3924	21	AAZ88974	Human MDR-3 DNA.
42	3481	56.5	3924	24	ABN95801	Gene #2299 used to
43	3453	56.1	4254	24	ABK63517	Rat sequence diffe
44	3405	55.3	3912	24	ABK63653	Rat sequence diffe
45	2844.5	46.2	4776	21	AAZ94744	Human ATP binding

ALIGNMENTS

```
RESULT 1
ABK83223
ID   ABK83223 standard; cDNA; 3699 BP.
XX
AC   ABK83223;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human transporter and ion channel, TRICH14, Incyte ID 7472030CB1, cDNA.
XX
KW   Human; ss; gene; transporter and ion channel; TRICH; transport disorder;
KW   neurological disorder; muscle disorder; immunological disorder; cancer;
KW   scleroderma; systemic lupus erythematosus; allergy; leukaemia;
KW   cell proliferative disorder; cervical cancer; breast cancer;
KW   neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;
KW   myotonic dystrophy; catatonia; endocrine disorder; diabetes;
KW   Grave's disease; gastrointestinal disorder; Crohn's disease;
KW   renal disorder; Good pasture's syndrome; viral infection; cirrhosis;
```

```
KW   bacterial infection; fungal infection; parasitic infection;
KW   protozoal infection; helminthic infection; cardiovascular disorder;
KW   atherosclerosis; hepatic disease.
XX
OS   Homo sapiens.
XX
PN   WO200240541-A2.
XX
PD   23-MAY-2002.
XX
PF   25-OCT-2001; 2001WO-US46055.
XX
PR   27-OCT-2000; 2000US-243989P.
PR   03-NOV-2000; 2000US-245904P.
PR   09-NOV-2000; 2000US-247673P.
PR   17-NOV-2000; 2000US-249661P.
PR   20-NOV-2000; 2000US-252232P.
PR   01-DEC-2000; 2000US-250790P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Tang YT,  Yue H,  Nguyen DB,  Hafalia AJA,  Elliott VS,  Lu Y;
PI   Walia NK,  Yao MG,  Baughn MR,  Gandhi AR,  Ding L,  Sanjanwala M;
PI   Ramkumar J,  Arvizu C,  Gietzen KJ,  Lal PG,  Azimzai Y,  Khan FA;
PI   Thangavelu K,  Thornton M,  Lu DAM,  Tribouley CM,  Warren BA;
PI   Ison CH,  Das D,  Raumann BE,  Policky JL,  Kearney L;
XX
DR   WPI; 2002-463570/49.
DR   P-PSDB; ABG61544.
XX
PT   New transporters and ion channels (TRICH) polypeptides, useful for
PT   diagnosing, preventing, and treating disorders associated with an
PT   abnormal expression or activity of TRICH, e.g. immunological, muscular
PT   or renal disorders  -
XX
PS   Claim 5; Page 172; 178pp; English.
XX
CC   The invention relates to human transporters and ion channels (TRICH)
CC   polypeptides, a naturally occurring amino acid sequence 90 % identical to
CC   TRICH, a biologically active fragment of TRICH or an immunogenic fragment
CC   of TRICH. Also included are an isolated polynucleotide encoding TRICH,
CC   a recombinant polynucleotide comprising a promoter sequence operably
CC   linked to the TRICH polynucleotide, a cell transformed with the
CC   recombinant polynucleotide, a transgenic organism comprising the
CC   recombinant polynucleotide, an isolated antibody that binds specifically
CC   to TRICH, and screening for compounds which bind to TRICH, modulate
CC   TRICH, modulate TRICH expression or are ant/agonists of TRICH.
CC   The polypeptides are useful for diagnosing, treating, and
CC   preventing transport, neurological, muscle, immunological disorders
CC   (e.g. scleroderma, systemic lupus erythematosus, allergies), cell
CC   proliferative disorders such as cancers (e.g. leukaemia, cervical or
CC   breast cancers), neurodegenerative disorders (e.g. Parkinson's disease,
CC   Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy,
CC   catatonia), endocrine disorders (e.g. diabetes, Grave's disease),
CC   gastrointestinal disorders (e.g. Crohn's disease), renal disorders
CC   (e.g. Good pasture's syndrome), viral, bacterial, fungal, parasitic,
CC   protozoal and helminthic infections, cardiovascular disorders (e.g.
CC   atherosclerosis), or hepatic diseases (e.g. cirrhosis) and many
CC   other diseases and disorders detailed in the specification. They can also
CC   be used in assessing the effects of exogenous compounds on the
CC   expression of nucleic acid and amino acid sequences of transporters and
CC   ion channels. TRICH or its fragments may also be used in screening for
CC   compounds that specifically bind to and modulate the activity of TRICH.
CC   The polynucleotides can be used to create knock-in humanised animals or
CC   transgenic animals to model human disease. The present sequence
CC   encodes a TRICH protein.
XX
SQ   Sequence 3699 BP; 1116 A; 707 C; 860 G; 1016 T; 0 other;

Alignment Scores:
Pred. No.:              0             Length:         3699
Score:                  5737.00       Matches:        1156
Percent Similarity:     93.78%        Conservative:   4
```

```
Best Local Similarity:  93.45%         Mismatches:    9
Query Match:            93.15%         Indels:        68
DB:                     24             Gaps:          4

US-09-873-409-5 (1-1222) x ABK83223 (1-3699)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
             ||||||||||||||||    |||||   |||||||||||||||||||||||||||
Db     145 ATGATCCTGGGTATACTGACATCACTGTTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 204

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
             :::||||||||||||||||||||||||||||||||||||||||   |||||||||
Db     205 TGTATAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCACACTAACACAACA 264

Qy      41 SerPheArg-----------------------------------------LeuThr 46
           :::::                                            |||||
Db     265 AATTATCAGAACTGTACTCAGTCTCAAGAGAAGCTGAATGAAGATATGACTCTGTTGACC 324

Qy      47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     325 CTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGTTACATACAGATTTCCTTG 384

Qy      67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     385 TGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAACAGTTTTTTCATTCAGTT 444

Qy      87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     445 TTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGTGAACTTAACACTCGCATG 504

Qy     107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
           |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     505 ACAGATGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATTGCTCTGTTGTTTCAA 564

Qy     126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     565 AACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAGGGCTGGAAACTCACC 624

Qy     146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     625 CTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCAGCATGTTCTAGGATG 684

Qy     166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     685 GTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCAGAA 744

Qy     186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     745 GAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAGGAGAAAGAACTTCAA 804

Qy     206 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeu 225

Db     804 ------------------------------------------------------------ 804

Qy     226 SerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 245
                                 ||||||||||||||||||||||||||||||||||||||||
Db     805 ---------------------AGGTATACACAGAATCTCAAAGATGCAAAGGATTTTGGC 843

Qy     246 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 265
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     844 ATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAATGGA 903

Qy     266 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 285
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     904 ACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGATAT 963

Qy     286 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 305
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     964 ACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGTAGTTATTGCATTGGA 1023

Qy     306 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 325
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1024 GCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTC 1083
```

```
Qy     326 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 345
           |||||||||||||||||||||||||||   |||||||||||||||||||||||||||||||
Db    1084 CAGGTTATTGATAAGAAACCCAGTATAGGTAACTTTTCCACAGCTGGATATAAACCTGAA 1143

Qy     346 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 365
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1144 TCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGGCCATCT 1203

Qy     366 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 385
           |||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db    1204 ATCAAGATTCTGAAAGGTCTGAATCTCGGAATTAAGTCTGGAGAGACAGTCGCCTTGGTC 1263

Qy     386 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 405
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1264 GGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCG 1323

Qy     406 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 425
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1324 GATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTAT 1383

Qy     426 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 445
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1384 CGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAAC 1443

Qy     446 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 465
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1444 AATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAA 1503

Qy     466 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 485
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1504 GCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAA 1563

Qy     486 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 505
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1564 AAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTT 1623

Qy     506 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 525
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1624 CGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAG 1683

Qy     526 SerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 545
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1684 TCAGCTGTTCAAGCTGCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTAGCA 1743

Qy     546 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 565
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1744 CACCGACTTTCTACTATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATGCTG 1803

Qy     566 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 585
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1804 GCGGAGAAAGGAGCGCATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTTGTG 1863

Qy     586 MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 605
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1864 ATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAA 1923

Qy     606 ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIleAsp 625
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1924 AGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGAC 1983

Qy     626 LysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLys 645
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1984 AAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTAAAA 2043

Qy     646 IleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerVal 665
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2044 ATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCTGTT 2103

Qy     666 LeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPhe 685
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2104 CTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATGTTT 2163

Qy     686 GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheVal 705
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2164 GGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTCGTC 2223

Qy     706 IleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAla 725
           ||||||||||||||||||||||||||||||||||
Db    2224 ATTTGGGTGTTATTTGCTTTGTCAGTTATTTCATG------------------------- 2259

Qy     726 GlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAsp 745
                                                                  |||||
Db    2260 ------------------------------------------------------CAGGAT 2265

Qy     746 IleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIle 765
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2266 ATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCCATA 2325

Qy     766 AspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAla 785
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2326 GATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAATGCA 2385

Qy     786 ThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeu 805
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2386 ACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTCCTG 2445

Qy     806 IleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThr 825
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2446 ATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATGACT 2505

Qy     826 GlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAla 845
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2506 GGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCT 2565

Qy     846 LeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyr 865
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2566 TTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATGTAT 2625

Qy     866 GluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySer 885
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2626 GAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGC 2685

Qy     886 CysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGly 905
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2686 TGTTATGCATTCAGCCATGCCTTTATATATTTTGCCTATGCGGCAGGGTTTCGATTTGGA 2745

Qy     906 AlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAla 925
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2746 GCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACTGCA 2805

Qy     926 IleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLys 945
           |||||||||||||||||||||||||:::||||||||||||||||||||||||||||||||
Db    2806 ATTGCATATGGAGCTATGGCCATCGGAGAAACGCTCGTTTTGGCTCCTGAATATTCCAAA 2865

Qy     946 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSer 965
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2866 GCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGACAGC 2925

Qy     966 ArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 985
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2926 CGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAAGTC 2985

Qy     986 SerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIle 1005
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2986 TCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGTATT 3045

Qy    1006 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 1025
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3046 GAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCTGTT 3105

Qy    1026 GlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAsp 1045
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3106 CAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTGGAT 3165

Qy    1046 AlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 1065
```

```
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3166 GCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAGCCT 3225

Qy    1066 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValVal 1085
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3226 GTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTGGTG 3285

Qy    1086 ProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGly 1105
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3286 CCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAAGGT 3345

Qy    1106 LeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGln 1125
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3346 CTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGCCAG 3405

Qy    1126 LysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAsp 1145
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3406 AAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTGGAT 3465

Qy    1146 GluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLys 1165
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3466 GAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGATAAA 3525

Qy    1166 AlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla 1185
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3526 GCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCA 3585

Qy    1186 AspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1205
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3586 GATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAGCTC 3645

Qy    1206 LeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1222
             |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3646 CTGAGAAATCGAGACATATATTTAAGTTAGTGAATGCACAGTCAGTGCAG 3696

RESULT 2
AAD03506
ID   AAD03506 standard; cDNA; 4279 BP.
XX
AC   AAD03506;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) allelic variant (Genotype D) cDNA.
XX
KW   Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KW   MDR1; drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag=  a
FT                   /product= "Dog P-glycoprotein (PGP) allelic variant
FT                   (Genotype D) protein"
FT   allele          replace (91, T)
FT                   /*tag=  b
FT   allele          replace (607, C)
FT                   /*tag=  c
FT   allele          replace (1001, T)
FT                   /*tag=  c
FT   allele          replace (3458, A)
FT                   /*tag=  c
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00310.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 9; Page 102-107; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) allelic variant
CC   (Genotype D) cDNA. The PGP enzyme functions as an efflux pump
CC   exporting small molecules across the cell membrane. This enzyme
CC   is a member of the ABC transporter family.
XX
SQ   Sequence 4279 BP; 1296 A; 833 C; 1009 G; 1141 T; 0 other;

Alignment Scores:
Pred. No.:              0              Length:        4279
Score:                  3533.50        Matches:       683
Percent Similarity:     73.13%         Conservative:  237
Best Local Similarity:  54.29%         Mismatches:    269
Query Match:            57.37%         Indels:        69
DB:                     22             Gaps:          9

US-09-873-409-5 (1-1222) x AAD03506 (1-4279)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
             |||:::::|||   :::|||::::::::::::||||||  |||||||||||    |||
Db     167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
             |||   |||   |||:::|||:::      :::        :::
Db     227 GTTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy      35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                                :::      ||||||     |||
Db     287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy      45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                           ||||||   |||||||||          |||:::
Db     347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy      60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
                |||||||||:::|||  |||    :::   |||   ||||||      :::||||||
Db     407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy      80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
             |||||||||||||||:::::::::   |||:::||||||||||||||||   |||:::
Db     467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy     100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
             |||||||||||||||||:::|||   |||:::  ||||||::::::|||||||||||||
Db     527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy     119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
             ||||||   :::   ||||||:::::::::||||||        |||      ||||||
Db     587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646
```

```
Qy     139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db     647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy     159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db     707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy     179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db     767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy     199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
Db     827 GGACAAAAGAAAGAACTTGAA--------------------------------------- 847

Qy     219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
Db     848 -----------------------------------------AGGTACAACAAAAATTTA 865

Qy     239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
Db     866 GAAGAAGCTAAAGGAATTGGGATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCC 925

Qy     259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
Db     926 GCTTTCTTATTGATCTATGCATCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTC 985

Qy     279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
Db     986 CTCTCCAGTGAA------TATACTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTA 1039

Qy     299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
Db    1040 ATTGGGGCTTTTAGTATTGGACAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGA 1099

Qy     319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
Db    1100 GGAGCAGCTTATGAAATCTTCAAGATAATTGACAATAAACCAAGCATTGACAGCTATTCG 1159

Qy     339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
Db    1160 AAGAGTGGACATAAACCAGATAATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTC 1219

Qy     359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
Db    1220 AGTTACCCTTCTCGAAAAGAAGTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGT 1279

Qy     379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
Db    1280 GGGCAGACAGTGGCGCTGGTTGGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTG 1339

Qy     399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
Db    1340 ATGCAGAGGCTCTATGACCCCACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGG 1399

Qy     419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db    1400 ACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTG 1459

Qy     439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
Db    1460 TTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAG 1519

Qy     459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
Db    1520 ATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTTATCATGAAACTACCTAATAAA 1579

Qy     479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db    1580 TTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATC 1639

Qy     499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db    1640 GCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCA 1699

Qy     519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 538
Db    1700 GCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGC 1759

Qy     539 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558
Db    1760 CGGACTACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCCGATGTCATTGCT 1819

Qy     559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578
Db    1820 GGTTTTGATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAG 1879

Qy     579 GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 598
Db    1880 GGCATTTACTTCAAACTTGTCACAATGCAG---ACAAGAGGAAATGAAATTGAGTTAGAA 1936

Qy     599 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 618
Db    1937 AATGCCACTGGTGAATCCAAAAGTGAAAGTGATGCCTTGGAAATG---TCTCCAAAAGAT 1993

Qy     619 IleLysSerAspPheIle------------------------------------------ 624
Db    1994 TCAGGGTCCAGTTTAATAAAAAGAAGATCAACTCGCAGGAGTATACATGCACCACAAGGC 2053

Qy     625 ---AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643
Db    2054 CAAGACAGAAAGCTTGGTACAAAAGAGGACTTGAATGAGAATGTACCTCCAGTTTCCTTC 2113

Qy     644 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 663
Db    2114 TGGAGGATTCTGAAGCTGAACTCAACTGAATGGCCTTATTTTGTGGTTGGTATATTTTGT 2173

Qy     664 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 683
Db    2174 GCTATTATAAACGGAGGCCTGCAACCAGCATTTTCAATAATATTTTCAAGGATTATAGGG 2233

Qy     684 MetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 702
Db    2234 ATCTTTACCCGAGATGAGGATCCTGAAACAAAACGACAGAATAGTAACATGTTTTCTGTA 2293

Qy     703 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722
Db    2294 TTGTTTCTAGTCCTTGGAATTATTTCTTTTATTACATTTTTCCTCCAGGGCTTCACATTT 2353

Qy     723 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 742
Db    2354 GGCAAAGCTGGGGAGATCCTCACTAAGCGGCTTCGATACATGGTTTTCAGATCCATGCTG 2413

Qy     743 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIle 762
Db    2414 AGACAGGATGTCAGCTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACAACCAGG 2473

Qy     763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 782
Db    2474 CTTGCCAATGATGCGGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTCATTACC 2533

Qy     783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 802
Db    2534 CAGAATATAGCAAATCTTGGGACAGGCATTATTATATCCTTAATCTATGGTTGGCAATTA 2593

Qy     803 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 822
Db    2594 ACACTTTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAA 2653

Qy     823 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 842
Db    2654 ATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCT 2713

Qy     843 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 862
```

```
Db   2714 ACAGAAGCCATCGAAAACTTCCGAACTGTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAA 2773

Qy    863 GlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 882
Db   2774 TACATGTATGCACAGAGTTTGCAAGTACCATACAGAAACTCTTTGAGGAAAGCACACATC 2833

Qy    883 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 902
Db   2834 TTCGGGGTCTCATTTTCTATCACCCAGGCAATGATGTATTTTTCCTATGCTGGCTGTTTC 2893

Qy    903 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 922
Db   2894 CGGTTTGGTGCCTACTTGGTGGCAAATGAGTTCATGAACTTTCAGGATGTTCTTTTGGTA 2953

Qy    923 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 942
Db   2954 TTCTCAGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAGGTCAGTTCATTTGCTCCTGAC 3013

Qy    943 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 962
Db   3014 TATGCCAAAGCCAAAGTATCAGCAGCCCACGTCATCATGATCATTGAAAAAAGCCCTCTG 3073

Qy    963 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 982
Db   3074 ATTGACAGCTACAGCCCTCACGGCCTCAAGCCAAATACGTTGGAAGGAAATGTGACATTT 3133

Qy    983 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002
Db   3134 AATGAGGTCGTGTTCAACTATCCCACTCGACCAGACATCCCCGTGCTCCAGGGGCTGAGC 3193

Qy   1003 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 1022
Db   3194 CTCGAGGTGAAGAAGGGCCAGACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGC 3253

Qy   1023 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1042
Db   3254 ACAGTTGTTCAGCTCCTAGAGCGCTTCTATGACCCCTTGGCTGGTTCAGTGCTAATTGAT 3313

Qy   1043 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 1062
Db   3314 GGCAAAGAGATAAAGCACCTGAATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCT 3373

Qy   1063 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082
Db   3374 CAGGAGCCCATCCTGTTTGACTGCAGCATTGCCGAGAACATTGCCTATGGAGACAACAGC 3433

Qy   1083 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1102
Db   3434 CGGGTCGTATCACATGAAGAGATTGTGCAGGCAGCCAAGGAGGCCAACATACACCACTTC 3493

Qy   1103 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122
Db   3494 ATCGAGACACTCCCTGAGAAATACAACACCAGAGTAGGAGACAAAGGAACCCAGCTCTCT 3553

Qy   1123 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 1142
Db   3554 GGTGGCCAGAAACAGCGCATTGCCATAGCTCGCGCTCTTGTTAGACAGCCTCATATTTTG 3613

Qy   1143 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 1162
Db   3614 CTTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCC 3673

Qy   1163 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 1182
Db   3674 CTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATCGCCCACCGCTTGTCCACCATC 3733

Qy   1183 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 1202
Db   3734 CAGAATGCAGATTTAATAGTGGTGTTTCAGAATGGCAAAGTCAAGGAGCATGGCACACAT 3793

Qy   1203 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220

Db   3794 CAACAGCTGCTGGCCCAGAAAGGCATCTATTTTTCCATGGTCAGTGTCCAGGCT 3847
```

```
RESULT 3
AAF86128
ID   AAF86128 standard; cDNA; 4195 BP.
XX
AC   AAF86128;
XX
DT   25-JUN-2001  (first entry)
XX
DE   Cynomologous monkey P-glycoprotein cDNA variant 1.
XX
KW   Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   efflux pump; ss.
XX
OS   Macaca fascicularis.
XX
FH   Key             Location/Qualifiers
FT   CDS             100..3951
FT                   /*tag=  a
FT                   /product= "PGP"
FT                   /note= "P-glycoprotein"
FT   misc_feature    376..384
FT                   /*tag=  b
FT                   /note= "Insertion of 9 nucleotides relative to PGP
FT                          allelic variant AAF86127"
XX
PN   WO200123565-A1.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26592.
XX
PR   28-SEP-1999;   99US-0156921.
PR   12-OCT-1999;   99US-0158818.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-Crespi DT,  Crespi CL;
XX
DR   WPI; 2001-316136/33.
DR   P-PSDB; AAB81065.
XX
PT   Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein
PT   (PGP) and homologous PGP polypeptides are useful for predicting
PT   bioavailability of compound and increasing PGP transporter activity in
PT   cell  -
XX
PS   Example 1; Page 59-65; 84pp; English.
XX
CC   This invention relates to a polynucleotide sequence encoding a
CC   cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the
CC   PGP protein. PGP, also known as multidrug transporter, MDR1 is a member
CC   of the ABC transporter superfamily. The enzyme serves as an efflux pump
CC   exporting small molecules across the cell membrane. The invention
CC   includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence
CC   and protein, and also that of an allelic variant. The PGP polynucleotide
CC   sequence is useful for increasing PGP transporter activity in a cell.
CC   Antisense sequences of the cDNA are useful for inhibiting PGP transport
CC   activity in a mammalian cell. They may also be used for increasing the
CC   bioavailability of a drug. The present sequence represents cDNA encoding
CC   cynomologous monkey P-glycoprotein. This sequence contains a 9 nucleotide
CC   insert compared to the PGP allelic variant given in AAF86127.
XX
SQ   Sequence 4195 BP; 1230 A; 801 C; 1039 G; 1125 T; 0 other;

Alignment Scores:
Pred. No.:              0              Length:        4195
Score:                  3533.00        Matches:       680
Percent Similarity:     73.55%         Conservative:  246
Best Local Similarity:  54.01%         Mismatches:    263
Query Match:            57.36%         Indels:        70
DB:                     22             Gaps:          10
```

```
US-09-873-409-5 (1-1222) x AAF86128 (1-4195)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db      250 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGAGCTGGACTTCCTCTCATGATGCTG 309

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIleSer---------------------GlyCys 33
Db      310 GTGTTTGGAGACATGACGGATACCTTTGCAAATGCAGGAAATTTAGGAGATTTAGGAGCT 369

Qy       34 LeuValGlnThrAsnThrTyrSer------------------------------------ 41
Db      370 CTGTTGTTTAACAACACTAATAGCAGTAATATCACTGATACAGTGCCCGTCATGAATCTG 429

Qy       42 ---------PhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIle 58
Db      430 GAGGAAGATATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTT 489

Qy       59 PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIle 78
Db      490 GCTGCTTACATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATT 549

Qy       79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp 98
Db      550 AGAAAACAGTTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGAT 609

Qy       99 IleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly 117
Db      610 GTTGGGGAGCTTAACACCCGGCTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGT 669

Qy      118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
Db      670 GACAAAATTGGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGA 729

Qy      138 LeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
Db      730 TTTACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTG 789

Qy      158 SerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr 177
Db      790 TCAGCTGCAGTCTGGGCAAAGATACTGTCTTCATTTACTGATAAAGAACTCTTAGCTTAT 849

Qy      178 SerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197
Db      850 GCAAAAGCTGGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTT 909

Qy      198 ArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrp 217
Db      910 GGAGGACAAAAGAAAGAACTCGAA------------------------------------ 933

Qy      218 PheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsn 237
Db      934 ---------------------------------------------AGGTACAACAAAAAT 948

Qy      238 LeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGly 257
Db      949 TTAGAAGAAGCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGT 1008

Qy      258 AlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeu 277
Db     1009 GCTGCTTTCCTGCTTATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTG 1068

Qy      278 IleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerVal 297
Db     1069 GTCCTCTCAAAGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTA 1122

Qy      298 IleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAla 317
Db     1123 TTAATTGGGGCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCA 1182

Qy      318 ArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPhe 337
Db     1183 AGAGGAGCAGCTTTTGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTAT 1242

Qy      338 SerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSer 357
Db     1243 TCGAAGAGTGGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCAC 1302

Qy      358 PheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLys 377
Db     1303 TTCAGTTACCCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAG 1362

Qy      378 SerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGln 397
Db     1363 AGTGGGCAGACGGTGGCCCTGGTTGGAAACAGCGGCTGTGGGAAGAGCACAACGGTCCAG 1422

Qy      398 LeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIle 417
Db     1423 CTGATGCAGAGGCTTTATGACCCCACAGAGGGCATGGTCAGTGTTGATGGACAGGATATT 1482

Qy      418 ArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProVal 437
Db     1483 AGGACCATAAACGTAAGGTTTCTACGGGAAATCATCGGTGTGGTGAGTCAGGAACCTGTA 1542

Qy      438 LeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGlu 457
Db     1543 TTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGTCGTGAAGATGTCACCATGGAT 1602

Qy      458 GluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsn 477
Db     1603 GAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCAG 1662

Qy      478 LysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArg 497
Db     1663 AAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGCTGAGTGGTGGGCAGAAGCAGAGG 1722

Qy      498 IleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThr 517
Db     1723 ATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGACGAGGCCACG 1782

Qy      518 SerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLys 537
Db     1783 TCAGCCTTGGACACAGAAAGTGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAA 1842

Qy      538 GlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIle 557
Db     1843 GGTCGGACCACCATTGTGATAGCTCATCGTTTGTCTACGGTTCGTAATGCCGACGTCATC 1902

Qy      558 ValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLys 577
Db     1903 GCTGGTTTCGATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAGCTCATGAAAGAG 1962

Qy      578 ArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle--------------------- 590
Db     1963 AAAGGCATTTACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAATTGAATTAGAA 2022

Qy      591 LysLysAlaAspGlu---------GlnMetGluSerMetThrTyrSerThrGluArgLys 607
Db     2023 AATGCAGCTGATGAATCCAAAAGTGAAATTGATACCTTGGAAATGTCTTCACATGATTCA 2082

Qy      608 ThrAsnSerLeu---------------ProLeuHisSerValLysSerIleLysSerAsp 622
Db     2083 GGATCCAGTCTAATAAGAAAAAGATCCACTCGTAGGAGTGTCCGTGGATCACAAGGCCAA 2142

Qy      623 PheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSer 642
Db     2143 ------GACAGAAAGCTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCC 2196

Qy      643 LeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeu 662
Db     2197 TTTTGGAGGATTATGAAGCTAAATTTAACTGAGTGGCCTTATTTTGTTGTTGGTGTATTT 2256

Qy      663 AlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIle 682
Db     2257 TGTGCCATTATAAATGGAGGTCTGCAACCAGCATTTGCAGTAATATTTTCAAAGATTATA 2316
```

```
Qy    683 ThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSer 701
Db   2317 GGGATTTTTACAAGAAATGATGATGCCGAAACAAAACGACAGAATAGTAACTTGTTTTCA 2376

Qy    702 MetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPhe 721
Db   2377 CTATTGTTTCTAGTCCTTGGAATTGTTTCTTTTATTACATTTTTCCTTCAGGGCTTCACA 2436

Qy    722 TyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMet 741
Db   2437 TTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATG 2496

Qy    742 LeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThr 761
Db   2497 CTCAGACAGGATGTGAGCTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACC 2556

Qy    762 IleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeu 781
Db   2557 AGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTATAATT 2616

Qy    782 ThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGlu 801
Db   2617 ACCCAGAATATAGCAAATCTTGGGACAGGAATAATTATATCCTTAATCTATGGTTGGCAA 2676

Qy    802 MetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThr 821
Db   2677 CTGACACTGTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATG 2736

Qy    822 AlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIle 841
Db   2737 AAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATC 2796

Qy    842 AlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPhe 861
Db   2797 GCTACTGAAGCAATAGAAAACTTCCGAACTGTTGTTTCTTTGACTCAGGAGCAGAAGTTT 2856

Qy    862 GluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGln 881
Db   2857 GAACATATGTATGATCAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACAC 2916

Qy    882 IleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGly 901
Db   2917 ATCTTTGGAATCACGTTTTCCTTCACGCAGGCAATGATGTATTTTTCCTATGCTGGATGT 2976

Qy    902 PheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIle 921
Db   2977 TTCCGGTTTGGAGCCTACTTGGTGGCACATAGTCTCATGAGCTTTGAGGATGTTCTGTTA 3036

Qy    922 ValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaPro 941
Db   3037 GTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCT 3096

Qy    942 GluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysPro 961
Db   3097 GACTATGCCAAAGCCAAAGTATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCT 3156

Qy    962 AsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGlu 981
Db   3157 TTGATTGACAGCTACAGCACAGAAGGCCTAAAGCCGAACACATTGGAAGGAAATGTCACA 3216

Qy    982 PheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeu 1001
Db   3217 TTTAATGAAGTTGTATTCAACTATCCCACCCGACTGGACATCCCAGTGCTTCAGGGGCTG 3276

Qy   1002 SerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLys 1021
Db   3277 AGCCTGGAAGTGAAGAAGGGCCAGACGCTGGCCCTGGTGGGCAGCAGTGGCTGTGGGAAG 3336

Qy   1022 SerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPhe 1041
Db   3337 AGCACGGTGGTCCAGCTCCTGGAGCGGTTCTATGACCCCTTGGCGGGGAAAGTGCTGCTT 3396

Qy   1042 AspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleVal 1061
Db   3397 GACGGCAAAGAAATAAAGCAACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTG 3456

Qy   1062 ProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsn 1081
Db   3457 TCCCAGGAGCCCATCCTGTTTGACTGCAGCATTAGTGAGAACATTGCCTATGGAGACAAC 3516

Qy   1082 SerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSer 1101
Db   3517 AGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCAGCCAAGGAGGCCAATATACACGCC 3576

Qy   1102 PheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeu 1121
Db   3577 TTCATCGAGTCACTGCCTAATAAATATAGCACCAGAGTAGGAGACAAAGGAACTCAGCTC 3636

Qy   1122 SerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIle 1141
Db   3637 TCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATT 3696

Qy   1142 LeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHis 1161
Db   3697 TTGCTTTTGGATGAAGCCACATCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAA 3756

Qy   1162 AlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAla 1181
Db   3757 GCCCTGGACAAAGCCAGAGAAGGCCGTACCTGCATTGTGATTGCTCACCGCCTGTCCACC 3816

Qy   1182 IleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThr 1201
Db   3817 ATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCACGGCACA 3876

Qy   1202 HisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db   3877 CATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3933
```

```
RESULT 4
AAZ49332
ID   AAZ49332 standard; cDNA; 3860 BP.
XX
AC   AAZ49332;
XX
DT   14-MAR-2000  (first entry)
XX
DE   Human wild-type multidrug resistance-1 (MDR-1) cDNA.
XX
KW   Multidrug resistance; MDR-1; P-glycoprotein;
KW   transmembrane efflux pump; haematopoietic stem cell; transduction;
KW   bone marrow transplantation; chemotherapy; radiation therapy; cancer;
KW   gene therapy; gene replacement; genetic defect; thalassaemia;
KW   Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
KW   cytokine; wild-type; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3843
FT                   /*tag= a
FT                   /product= "Human wild-type MDR-1 protein"
FT   mutation        replace (553..555, GTT)
FT                   /*tag= b
FT                   /note= "cDNA sequence of G185V human mutant MDR-1 given
FT                          in AAZ49333"
XX
PN   WO9961589-A2.
XX
PD   02-DEC-1999.
XX
PF   27-MAY-1999;   99WO-US11825.
XX
PR   28-MAY-1998;   98US-0086988.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
```

```
XX
PI   Sorrentino B,  Bunting K;
XX
DR   WPI; 2000-072615/06.
DR   P-PSDB; AAY58186.
XX
PT   Ex vivo expansion of hematopoietic stem cells transduced with a
PT   sequence encoding human multidrug resistance-1, used for bone marrow
PT   transplantation -
XX
PS   Claim 10; Page 68-70; 113pp; English.
XX
CC   This sequence represents cDNA encoding human wild-type
CC   multidrug resistance protein MDR-1. MDR-1 is a transmembrane
CC   efflux pump, responsible for the export of drugs from cells,
CC   particularly cancer cells. Wild-type MDR-1 shows increased
CC   resistance to etoposide and decreased resistance to vinca
CC   alkaloids compared with a mutant form (AAY58187) where the Gly at
CC   position 185 is replaced by Val. The invention relates to transducing
CC   haematopoietic stem cells with nucleic acid encoding an MDR protein
CC   and culturing the modified cells. The modified haematopoietic stem
CC   cells are useful in bone marrow transplantation (to reconstitute
CC   haematopoietic systems in patients who have undergone chemotherapy or
CC   radiation therapy) and in ex vivo gene therapy of genetic defects in
CC   cells derived from haematopoietic stem cells, e.g., thalassaemia,
CC   Gauchier's disease, sickle cell anaemia or leukaemia. The modified
CC   cells can also be used to identify factors involved in regulating
CC   proliferation and differentiation in haematopoietic stem cells.
CC   Haematopoietic stem cells that express MDR-1 will be protected against
CC   chemotherapeutic agents, so can be engrafted while the patient is
CC   undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells
CC   provides sufficient cells to permit standard biochemical analysis.
CC   Overexpression of MDR-1 allows cytokine-driven expansion of
CC   haematopoietic stem cells by at least 10-fold compared with a maximum
CC   of 4-fold in known procedures.
XX
SQ   Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;

Alignment Scores:
Pred. No.:              0          Length:         3860
Score:                  3530.50    Matches:        682
Percent Similarity:     74.00%     Conservative:   246
Best Local Similarity:  54.39%     Mismatches:     263
Query Match:            57.32%     Indels:         63
DB:                     21         Gaps:           10

US-09-873-409-5 (1-1222) x AAZ49332 (1-3860)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            |||::::::|||   |||||:::::::::::|||||   |||||||||||    |||
Db      151 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210

Qy       21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
            |||   ||||||||:::|||                           :::|||:::|||
Db      211 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 270

Qy       32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                  ::::::        |||        ||||||
Db      271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330

Qy       44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                  |||      ||||||   ||||||||          |||:::         |||
Db      331 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 390

Qy       62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
            ||||||:::|||   |||   :::   |||   ||||||         :::|||||||||||
Db      391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450

Qy       82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
            ||||||||:::::::::   |||:::||||||||||||||         |||:::||||||
Db      451 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 510

Qy      102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
            ||||||||||||:::|||   |||:::   |||:::   ||||||:::::::::||||||||||||||||
Db      511 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGGTATTGGTGACAAAATT 570

Qy      121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
               :::   ||||||:::|||:::||||||         |||      ||||||      :::
Db      571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 630

Qy      141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
            ||||||||||||||||||||   |||:::   ||||||::::::      |||||||||
Db      631 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

Qy      161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
                  ::::::::::::      |||   |||   |||||||||   ||||||:::||||||
Db      691 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy      181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
            |||||||||||||||||||||||::::::||||||||||||||||||         |||
Db      751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810

Qy      201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
            :::|||||||||:::
Db      811 AAGAAAGAACTTGAA--------------------------------------------- 825

Qy      221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                               ||||||   :::||||||:::::
Db      826 ------------------------------------AGGTACAACAAAAATTTAGAAGAA 849

Qy      241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
            ||||||      |||||||||:::   |||   :::   :::|||:::||||||      :::
Db      850 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 909

Qy      261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
               :::      :::|||   |||||||||||||||||||:::|||:::|||:::
Db      910 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 969

Qy      281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
            ||||||      |||:::||||||   ||||||   |||||||||||||||:::
Db      970 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1023

Qy      301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
            ::::::   :::|||   |||   |||   ||||||   ||||||||||||
Db     1024 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1083

Qy      321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
            |||:::   ||||||::::::||||||   ||||||||||||||:::::::|||      :::
Db     1084 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1143

Qy      341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
            |||:::|||||||::::::|||:::|||      :::||||||:::||||||   |||:::|||
Db     1144 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1203

Qy      361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
            ||||||||      :::||||||||||||||||||||||||::::::::::|||||:::
Db     1204 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1263

Qy      381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
            ||||||||||||||||||      :::|||   ||||||||||||      |||||||||:::|||
Db     1264 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1323

Qy      401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
            ||||||||||||||      :::|||   :::      |||||||||      :::
Db     1324 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1383

Qy      421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
            ||||||||      |||:::   |||||||||||||||||||||||||||||||||
Db     1384 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1443

Qy      441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
            |||||||||:::   ||||||:::|||||||||:::::::|||||      :::|||:::|||
Db     1444 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1503

Qy      461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
            :::|||   :::||||||||||||||||||||||||:::      |||:::||||||:::
```

```
Db   1504 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1563

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500

Db   1564 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 1623

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520

Db   1624 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 1683

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540

Db   1684 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACC 1743

Qy    541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560

Db   1744 ACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC 1803

Qy    561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580

Db   1804 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATT 1863

Qy    581 TyrTyrSerLeuValMetSerGln------------------AspIleLysLysAlaAsp 594

Db   1864 TACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 1923

Qy    595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614

Db   1924 GATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGT 1983

Qy    615 SerVal------------------LysSerIleLys---SerAspPheIleAspLysAla 627

Db   1984 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAG 2043

Qy    628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647

Db   2044 CTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATG 2103

Qy    648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667

Db   2104 AAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAAT 2163

Qy    668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687

Db   2164 GGAGGCCTGCAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGA 2223

Qy    688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706

Db   2224 ATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCC 2283

Qy    707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726

Db   2284 CTTGGAATTATTTCTTTTATTACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGA 2343

Qy    727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746

Db   2344 GAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTG 2403

Qy    747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766

Db   2404 AGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGAT 2463

Qy    767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786

Db   2464 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCA 2523

Qy    787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806

Db   2524 AATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTC 2583

Qy    807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826

Db   2584 TTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGA 2643

Qy    827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846

Db   2644 CAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 2703

Qy    847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866

Db   2704 GAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCT 2763

Qy    867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886

Db   2764 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 2823

Qy    887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906

Db   2824 TTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCC 2883

Qy    907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926

Db   2884 TACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTT 2943

Qy    927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946

Db   2944 GTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCC 3003

Qy    947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966

Db   3004 AAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTAC 3063

Qy    967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986

Db   3064 AGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTA 3123

Qy    987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006

Db   3124 TTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAG 3183

Qy   1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026

Db   3184 AAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAG 3243

Qy   1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046

Db   3244 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATA 3303

Qy   1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066

Db   3304 AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATC 3363

Qy   1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086

Db   3364 CTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCA 3423

Qy   1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106

Db   3424 CAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTG 3483

Qy   1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126

Db   3484 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAA 3543

Qy   1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146

Db   3544 CAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAA 3603

Qy   1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166

Db   3604 GCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCC 3663

Qy   1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186

Db   3664 AGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGAC 3723
```

```
Qy     1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
            |||||||||||      |||||::::::||||||  |||||||||||:::||||||
Db     3724 TTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTG 3783

Qy     1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
                  :::  |||||||||   :::|||:::   |||:::
Db     3784 GCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3825

RESULT 5
ABA94365
ID   ABA94365 standard; DNA; 3860 BP.
XX
AC   ABA94365;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human BCRP DNA related seq Id No. 1.
XX
KW   Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;
KW   hepatotropic; neurodegenerative; cytostatic; antianemic; muscular; BCRP;
KW   cardiant; gene therapy; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3843
FT                   /*tag=  a
XX
PN   WO200192877-A2.
XX
PD   06-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US17459.
XX
PR   31-MAY-2000; 2000US-0584586.
PR   29-MAY-2001; 2001US-0866866.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI   Sorrentino B,  Schuetz J;
XX
DR   WPI; 2002-114368/15.
DR   P-PSDB; ABB07266.
XX
PT   Identifying a stem cell, for treating e.g., muscular dystrophy,
PT   myocardial infarction, Parkinson's disease, or neurodegenerative
PT   disorders, comprises detecting the expression of an ATP transport
PT   protein (BCRP) by a cell  -
XX
PS   Disclosure; Page 53-55; 87pp; English.
XX
CC   The invention provides a method of identifying and/or isolating a stem
CC   cell that involves detecting the expression of an ATP transport protein
CC   containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
CC   comprising stem cells. The isolated stem cells may be used in the
CC   treatment of diseases such as muscular dystrophy, degenerative liver
CC   disorder, myocardial infarction, Parkinson's disease, degenerative
CC   disorders of the brain, and for tissue regeneration or replacement.
CC   Haematopoietic cells can be used in bone marrow transplants (e.g., for
CC   treatment of leukemia) and for ex vivo gene therapy for treating blood
CC   diseases such as sickle cell anemia and thalassemia. The stem cells can
CC   also be used as cell targets in gene therapy protocols. The present
CC   sequence represents a sequence related to the BCRP for which no relevant
CC   information has been provided in the specification.
XX
SQ   Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;

Alignment Scores:
Pred. No.:              0          Length:          3860
Score:                  3530.50    Matches:         682
Percent Similarity:     74.00%     Conservative:    246
Best Local Similarity:  54.39%     Mismatches:      263
Query Match:            57.32%     Indels:          63
DB:                     24         Gaps:            10

US-09-873-409-5 (1-1222) x ABA94365 (1-3860)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            |||::::::|||   ||||||::::::::::::||||||   |||||||||||   |||
Db      151 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210

Qy       21 ValLeuGlyGluMetSerAsp-------------------------AsnLeuIleSer 31
            |||   ||||||||:::|||                           :::|||:::|||
Db      211 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 270

Qy       32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                           ::::::       |||       ||||||
Db      271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330

Qy       44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                  |||         ||||||  ||||||||             |||:::     |||
Db      331 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 390

Qy       62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
            ||||||:::|||   |||    :::   |||   ||||||      :::||||||||||||
Db      391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450

Qy       82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
            |||||||||::::::::::   |||:::|||||||||||||||      |||:::||||||
Db      451 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 510

Qy      102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
            ||||||||||||:::|||   |||:::   ||||||:::::::||||||||||||||||||
Db      511 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGGTATTGGTGACAAAATT 570

Qy      121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
               :::   ||||||:::|||:::||||||       |||     ||||||       :::
Db      571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 630

Qy      141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
            |||||||||||||||||||||   |||:::   ||||||::::::    ||||||||||
Db      631 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

Qy      161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
                ::::::::::::   |||   |||   |||||||||   ||||||:::||||||
Db      691 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy      181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
            ||||||||||||||||||||||||::::::||||||||||||||||||||       |||
Db      751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810

Qy      201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
            :::|||||||||:::
Db      811 AAGAAAGAACTTGAA--------------------------------------------- 825

Qy      221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                                 ||||||   :::||||||::::::
Db      826 ------------------------------------AGGTACAACAAAAATTTAGAAGAA 849

Qy      241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
            ||||||     |||||||||:::   |||   :::   :::|||:::||||||     :::
Db      850 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 909

Qy      261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
               :::       :::|||   ||||||||||||||||||||:::|||:::|||:::
Db      910 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 969

Qy      281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
            ||||||      |||:::||||||   ||||||   |||||||||||||||:::
Db      970 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1023

Qy      301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
            ::::::   :::|||   |||   |||         |||  ||||||   ||||||||||||
Db     1024 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1083

Qy      321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
```

```
Db   1084 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1143

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   1144 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1203

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   1204 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1263

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   1264 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1323

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   1324 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1383

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   1384 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1443

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   1444 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1503

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   1504 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1563

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   1564 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 1623

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   1624 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 1683

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db   1684 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACC 1743

Qy    541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db   1744 ACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC 1803

Qy    561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db   1804 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATT 1863

Qy    581 TyrTyrSerLeuValMetSerGln------------------AspIleLysLysAlaAsp 594
Db   1864 TACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 1923

Qy    595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
Db   1924 GATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGT 1983

Qy    615 SerVal------------------LysSerIleLys---SerAspPheIleAspLysAla 627
Db   1984 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAG 2043

Qy    628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
Db   2044 CTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATG 2103

Qy    648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
Db   2104 AAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAAT 2163

Qy    668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
Db   2164 GGAGGCCTGCAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGA 2223

Qy    688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db   2224 ATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCC 2283

Qy    707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db   2284 CTTGGAATTATTTCTTTTATTACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGA 2343

Qy    727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db   2344 GAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTG 2403

Qy    747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
Db   2404 AGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGAT 2463

Qy    767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db   2464 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCA 2523

Qy    787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
Db   2524 AATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTC 2583

Qy    807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db   2584 TTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGA 2643

Qy    827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db   2644 CAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 2703

Qy    847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db   2704 GAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCT 2763

Qy    867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db   2764 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 2823

Qy    887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db   2824 TTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCC 2883

Qy    907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db   2884 TACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTT 2943

Qy    927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db   2944 GTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCC 3003

Qy    947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db   3004 AAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTAC 3063

Qy    967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db   3064 AGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTA 3123

Qy    987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db   3124 TTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAG 3183

Qy   1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
Db   3184 AAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAG 3243

Qy   1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
Db   3244 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATA 3303
```

```
Qy    1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
             |||  ||||||||||||||||||||:::   :::  ||||||  |||||||||:::
Db    3304 AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATC 3363

Qy    1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
             ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3364 CTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCA 3423

Qy    1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
               :::||||||      ||||||      ||||||||||||:::|||||||||   |||
Db    3424 CAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTG 3483

Qy    1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
             |||   ||||||:::|||:::||||||   ||||||   |||||||||||||||||||||
Db    3484 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAA 3543

Qy    1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
             ||||||:::||||||||||||||||||||||:::::::::|||   ||||||||||||||||||
Db    3544 CAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTGCTTTTGGATGAA 3603

Qy    1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
             ||||||||||||||||||     :::||||||||||||||||||   ||||||||||||||
Db    3604 GCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCC 3663

Qy    1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
             |||   ||||||||||||:::|||:::   ||||||||||||   |||||||||||||||
Db    3664 AGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGAC 3723

Qy    1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
             ||||||||||||      ||||||:::::||||||   |||||||||||:::||||||
Db    3724 TTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTG 3783

Qy    1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
                :::   |||||||||   :::|||:::   |||:::
Db    3784 GCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3825
```

```
RESULT 6
AAD03504
ID   AAD03504 standard; cDNA; 4279 BP.
XX
AC   AAD03504;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) allelic variant (Genotype A) cDNA.
XX
KW   Dog; P-glycoprotein allelic variant; multidrug transporter; MDR1;
KW   drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag=  a
FT                   /product= "Dog P-glycoprotein (PGP) allelic variant
FT                   (Genotype A) protein"
FT   allele          replace (607, C)
FT                   /*tag=  b
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00308.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 9; Page 85-90; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) allelic variant
CC   (Genotype A) cDNA. The PGP enzyme functions as an efflux pump
CC   exporting small molecules across the cell membrane. This enzyme
CC   is a member of the ABC transporter family.
XX
SQ   Sequence 4279 BP; 1295 A; 833 C; 1008 G; 1143 T; 0 other;

Alignment Scores:
Pred. No.:              0              Length:         4279
Score:                  3530.50        Matches:        682
Percent Similarity:     73.13%         Conservative:   238
Best Local Similarity:  54.21%         Mismatches:     269
Query Match:            57.32%         Indels:         69
DB:                     22             Gaps:           9

US-09-873-409-5 (1-1222) x AAD03504 (1-4279)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
             |||::::::|||   :::|||::::::::::::||||||  ||||||||||||   |||
Db     167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
             |||   |||   |||:::|||:::      :::
Db     227 GTTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy      35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                                 :::      ||||||   |||
Db     287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy      45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                              ||||||   ||||||||||         |||:::
Db     347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy      60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
                |||||||||:::|||   |||   :::   |||   ||||||      :::||||||
Db     407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy      80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
             |||||||||||||||:::::::::   ||||:::|||||||||||||||      |||:::
Db     467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy     100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
             ||||||||||||||||||:::|||   |||:::   ||||||::::::||||||||||||
Db     527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy     119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
             ||||||   :::   |||||||:::::::::||||||      |||      ||||||
Db     587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy     139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
                :::||||||||||||||||||||||   |||:::   ||||||::::::      |||
Db     647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706
```

```
Qy    159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db    707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy    179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db    767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy    199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
Db    827 GGACAAAAGAAAGAACTTGAA--------------------------------------- 847

Qy    219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
Db    848 ---------------------------------------------AGGTACAACAAAAATTTA 865

Qy    239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
Db    866 GAAGAAGCTAAAGGAATTGGGATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCC 925

Qy    259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
Db    926 GCTTTCTTATTGATCTATGCATCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTC 985

Qy    279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
Db    986 CTCTCCAGTGAA------TATTCTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTA 1039

Qy    299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
Db   1040 ATTGGGGCTTTTAGTATTGGACAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGA 1099

Qy    319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
Db   1100 GGAGCAGCTTATGAAATCTTCAAGATAATTGACAATAAACCAAGCATTGACAGCTATTCG 1159

Qy    339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
Db   1160 AAGAGTGGACATAAACCAGATAATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTC 1219

Qy    359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
Db   1220 AGTTACCCTTCTCGAAAAGAAGTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGT 1279

Qy    379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
Db   1280 GGGCAGACAGTGGCGCTGGTTGGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTG 1339

Qy    399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
Db   1340 ATGCAGAGGCTCTATGACCCCACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGG 1399

Qy    419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db   1400 ACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTG 1459

Qy    439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
Db   1460 TTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAG 1519

Qy    459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
Db   1520 ATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTTATCATGAAACTACCTAATAAA 1579

Qy    479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db   1580 TTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATC 1639

Qy    499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db   1640 GCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCA 1699

Qy    519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 538
Db   1700 GCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGC 1759

Qy    539 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558
Db   1760 CGGACTACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCCGATGTCATTGCT 1819

Qy    559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578
Db   1820 GGTTTTGATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAG 1879

Qy    579 GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 598
Db   1880 GGCATTTACTTCAAACTTGTCACAATGCAG---ACAAGAGGAAATGAAATTGAGTTAGAA 1936

Qy    599 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 618
Db   1937 AATGCCACTGGTGAATCCAAAAGTGAAAGTGATGCCTTGGAAATG---TCTCCAAAAGAT 1993

Qy    619 IleLysSerAspPheIle------------------------------------------ 624
Db   1994 TCAGGGTCCAGTTTAATAAAAAGAAGATCAACTCGCAGGAGTATACATGCACCACAAGGC 2053

Qy    625 ---AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643
Db   2054 CAAGACAGAAAGCTTGGTACAAAAGAGGACTTGAATGAGAATGTACCTCCAGTTTCCTTC 2113

Qy    644 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 663
Db   2114 TGGAGGATTCTGAAGCTGAACTCAACTGAATGGCCTTATTTTGTGGTTGGTATATTTTGT 2173

Qy    664 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 683
Db   2174 GCTATTATAAACGGAGGCCTGCAACCAGCATTTTCAATAATATTTTCAAGGATTATAGGG 2233

Qy    684 MetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 702
Db   2234 ATCTTTACCCGAGATGAGGATCCTGAAACAAAACGACAGAATAGTAACATGTTTTCTGTA 2293

Qy    703 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722
Db   2294 TTGTTTCTAGTCCTTGGAATTATTTCTTTTATTACATTTTTCCTCCAGGGCTTCACATTT 2353

Qy    723 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 742
Db   2354 GGCAAAGCTGGGGAGATCCTCACTAAGCGGCTTCGATACATGGTTTTCAGATCCATGCTG 2413

Qy    743 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIle 762
Db   2414 AGACAGGATGTCAGCTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACAACCAGG 2473

Qy    763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 782
Db   2474 CTTGCCAATGATGCGGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTCATTACC 2533

Qy    783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 802
Db   2534 CAGAATATAGCAAATCTTGGGACAGGCATTATTATATCCTTAATCTATGGTTGGCAATTA 2593

Qy    803 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 822
Db   2594 ACACTTTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAA 2653

Qy    823 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 842
Db   2654 ATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCT 2713

Qy    843 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 862
Db   2714 ACAGAAGCCATCGAAAACTTCCGAACTGTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAA 2773

Qy    863 GlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 882
```

```
               ||||||    :::   ||||||      :::||||||:::    :::||||||   |||
Db      2774 TACATGTATGCACAGAGTTTGCAAGTACCATACAGAAACTCTTTGAGGAAAGCACACATC 2833

Qy       883 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 902
                |||     ::::::   :::     |||    :::||||||:::||||||      |||
Db      2834 TTCGGGGTCTCATTTTCTATCACCCAGGCAATGATGTATTTTTCCTATGCTGGCTGTTTC 2893

Qy       903 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 922
             ||||||||||||||||||:::                |||      :::    :::    :::|||
Db      2894 CGGTTTGGTGCCTACTTGGTGGCAAATGAGTTCATGAACTTTCAGGATGTTCTTTTGGTA 2953

Qy       923 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 942
             |||:::||||||      :::||||||||||||:::|||:::             ||||||:::
Db      2954 TTCTCAGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAGGTCAGTTCATTTGCTCCTGAC 3013

Qy       943 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 962
             |||:::|||||||||          |||||||||:::         ::::::||||||    |||
Db      3014 TATGCCAAAGCCAAAGTATCAGCAGCCCACGTCATCATGATCATTGAAAAAAGCCCTCTG 3073

Qy       963 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 982
             |||||||||   |||          |||    ||||||:::|||    |||||||||:::    |||
Db      3074 ATTGACAGCTACAGCCCTCACGGCCTCAAGCCAAATACGTTGGAAGGAAATGTGACATTT 3133

Qy       983 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002
                ||||||    |||    ||||||      |||||||||:::      :::|||:::|||||||||
Db      3134 AATGAGGTCGTGTTCAACTATCCCACTCGACCAGACATCCCCGTGCTCCAGGGGCTGAGC 3193

Qy      1003 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 1022
             |||    :::::::::|||:::|||:::|||    ||||||||||||||||||||||||
Db      3194 CTCGAGGTGAAGAAGGGCCAGACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGC 3253

Qy      1023 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1042
             |||    ||||||||||||:::|||      ||||||||||:::    |||      ||||||    |||
Db      3254 ACAGTTGTTCAGCTCCTAGAGCGCTTCTATGACCCCTTGGCTGGTTCAGTGCTAATTGAT 3313

Qy      1043 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 1062
             |||    :::    |||    |||||||||||||||||||||:::    :::    ||||||
Db      3314 GGCAAAGAGATAAAGCACCTGAATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCT 3373

Qy      1063 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082
             |||||||||:::|||||||||||||||||||||||||||||||||||||||||||||||
Db      3374 CAGGAGCCCATCCTGTTTGACTGCAGCATTGCCGAGAACATTGCCTATGGAGACAACAGC 3433

Qy      1083 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1102
             ||||||||       :::||||||    :::||||||       ||||||||||||    |||
Db      3434 CGGGTCGTATCACATGAAGAGATTATGCAGGCAGCCAAGGAGGCCAACATACACCACTTC 3493

Qy      1103 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122
             ||||||    ||||||||||||||||||||||:::||||||    ||||||    |||||||||
Db      3494 ATCGAGACACTCCCTGAGAAATACAACACCAGAGTAGGAGACAAAGGAACCCAGCTCTCT 3553

Qy      1123 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 1142
             ||||||||||||||||||:::||||||||||||||||||::::::::|||    ||||||
Db      3554 GGTGGCCAGAAACAGCGCATTGCCATAGCTCGCGCTCTTGTTAGACAGCCTCATATTTTG 3613

Qy      1143 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 1162
             |||||||||||||||||||||||||||       :::|||||||||||||||||    |||
Db      3614 CTTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCC 3673

Qy      1163 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 1182
             |||||||||||||||    ||||||||||||:::|||:::    ||||||||||||    |||
Db      3674 CTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATCGCCCACCGCTTGTCCACCATC 3733

Qy      1183 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 1202
             ||||||||||||||||||||||||        |||||||||:::||||||    |||||||||
Db      3734 CAGAATGCAGATTTAATAGTGGTGTTTCAGAATGGCAAAGTCAAGGAGCATGGCACACAT 3793

Qy      1203 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
             |||:::||||||      :::    |||||||||    :::|||:::    |||:::
Db      3794 CAACAGCTGCTGGCCCAGAAAGGCATCTATTTTTCCATGGTCAGTGTCCAGGCT 3847
```

RESULT 7
AAD03505

```
ID   AAD03505 standard; cDNA; 4279 BP.
XX
AC   AAD03505;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.
XX
KW   Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KW   MDR1; drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag= a
FT                   /product= "Dog P-glycoprotein (PGP) allelic variant
FT                   (Genotype B) protein"
FT   allele          replace (91, T)
FT                   /*tag= b
FT   allele          replace (607, C)
FT                   /*tag= c
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00309.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 9; Page 93-99; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) allelic variant
CC   (Genotype B) cDNA. The PGP enzyme functions as an efflux pump
CC   exporting small molecules across the cell membrane. This enzyme
CC   is a member of the ABC transporter family.
XX
SQ   Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;

Alignment Scores:
Pred. No.:               0              Length:        4279
Score:                   3530.50        Matches:       682
Percent Similarity:      73.13%         Conservative:  238
Best Local Similarity:   54.21%         Mismatches:    269
Query Match:             57.32%         Indels:        69
DB:                      22             Gaps:          9

US-09-873-409-5 (1-1222) x AAD03505 (1-4279)
```

```
Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
Db    227 GTTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy     35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
Db    287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy     45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db    347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy     60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db    407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy     80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db    467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy    100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db    527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy    119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db    587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy    139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db    647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy    159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db    707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy    179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db    767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy    199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
Db    827 GGACAAAAGAAAGAACTTGAA--------------------------------------- 847

Qy    219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
Db    848 -------------------------------------------AGGTACAACAAAAATTTA 865

Qy    239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
Db    866 GAAGAAGCTAAAGGAATTGGGATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCC 925

Qy    259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
Db    926 GCTTTCTTATTGATCTATGCATCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTC 985

Qy    279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
Db    986 CTCTCCAGTGAA------TATTCTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTA 1039

Qy    299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
Db   1040 ATTGGGGCTTTTAGTATTGGACAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGA 1099

Qy    319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
Db   1100 GGAGCAGCTTATGAAATCTTCAAGATAATTGACAATAAACCAAGCATTGACAGCTATTCG 1159

Qy    339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
Db   1160 AAGAGTGGACATAAACCAGATAATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTC 1219

Qy    359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
Db   1220 AGTTACCCTTCTCGAAAAGAAGTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGT 1279

Qy    379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
Db   1280 GGGCAGACAGTGGCGCTGGTTGGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTG 1339

Qy    399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
Db   1340 ATGCAGAGGCTCTATGACCCCACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGG 1399

Qy    419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db   1400 ACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTG 1459

Qy    439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
Db   1460 TTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAG 1519

Qy    459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
Db   1520 ATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTTATCATGAAACTACCTAATAAA 1579

Qy    479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db   1580 TTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATC 1639

Qy    499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db   1640 GCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCA 1699

Qy    519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 538
Db   1700 GCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGC 1759

Qy    539 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558
Db   1760 CGGACTACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCCGATGTCATTGCT 1819

Qy    559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578
Db   1820 GGTTTTGATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAG 1879

Qy    579 GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 598
Db   1880 GGCATTTACTTCAAACTTGTCACAATGCAG---ACAAGAGGAAATGAAATTGAGTTAGAA 1936

Qy    599 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 618
Db   1937 AATGCCACTGGTGAATCCAAAAGTGAAAGTGATGCCTTGGAAATG---TCTCCAAAAGAT 1993

Qy    619 IleLysSerAspPheIle------------------------------------------ 624
Db   1994 TCAGGGTCCAGTTTAATAAAAAGAAGATCAACTCGCAGGAGTATACATGCACCACAAGGC 2053

Qy    625 ---AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643
Db   2054 CAAGACAGAAAGCTTGGTACAAAAGAGGACTTGAATGAGAATGTACCTCCAGTTTCCTTC 2113

Qy    644 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 663
Db   2114 TGGAGGATTCTGAAGCTGAACTCAACTGAATGGCCTTATTTTGTGGTTGGTATATTTTGT 2173

Qy    664 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 683
Db   2174 GCTATTATAAACGGAGGCCTGCAACCAGCATTTTCAATAATATTTTCAAGGATTATAGGG 2233

Qy    684 MetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 702
```

```
              :::|||     :::   |||      |||   :::   ::::::   ::::::|||:::
Db     2234 ATCTTTACCCGAGATGAGGATCCTGAAACAAAACGACAGAATAGTAACATGTTTTCTGTA 2293

Qy      703 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722
              :::|||::::::|||||||:::|||   |||:::::::::|||:::|||||      :::
Db     2294 TTGTTTCTAGTCCTTGGAATTATTTCTTTTATTACATTTTTCCTCCAGGGCTTCACATTT 2353

Qy      723 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 742
            |||:::||||||||||||||||||   ||||||||||::::::   |||::::::||||||
Db     2354 GGCAAAGCTGGGGAGATCCTCACTAAGCGGCTTCGATACATGGTTTTCAGATCCATGCTG 2413

Qy      743 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIle 762
                  ||||||::::::|||||||||:::   :::|||:::||||||   |||||||||
Db     2414 AGACAGGATGTCAGCTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACAACCAGG 2473

Qy      763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 782
            ||||||   |||   ||||||::::::||||||   |||||||||:::   |||:::|||
Db     2474 CTTGCCAATGATGCGGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTCATTACC 2533

Qy      783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 802
            ||||||      |||:::|||         :::||||||||   ||||||||||||::::::
Db     2534 CAGAATATAGCAAATCTTGGGACAGGCATTATTATATCCTTAATCTATGGTTGGCAATTA 2593

Qy      803 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 822
            |||   |||:::|||:::|||   |||::::::|||   |||::::::|||
Db     2594 ACACTTTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAA 2653

Qy      823 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 842
               ::::::|||   |||   |||||||||:::||||||:::   ||||||||||||||
Db     2654 ATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCT 2713

Qy      843 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 862
            |||||||||:::||||||   ||||||:::||||||||||||||||||:::   ||||||
Db     2714 ACAGAAGCCATCGAAAACTTCCGAACTGTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAA 2773

Qy      863 GlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 882
                  ||||||   :::   ||||||      :::|||||||:::   :::||||||   |||
Db     2774 TACATGTATGCACAGAGTTTGCAAGTACCATACAGAAACTCTTGAGGAAAGCACACATC 2833

Qy      883 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 902
               |||      ::::::   :::   |||   :::||||||:::||||||   |||
Db     2834 TTCGGGGTCTCATTTTCTATCACCCAGGCAATGATGTATTTTTCCTATGCTGGCTGTTTC 2893

Qy      903 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 922
            ||||||||||||||||||:::               |||      :::   :::   :::|||
Db     2894 CGGTTTGGTGCCTACTTGGTGGCAAATGAGTTCATGAACTTTCAGGATGTTCTTTTGGTA 2953

Qy      923 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 942
            |||:::||||||   :::||||||||||||:::|||:::             ||||||:::
Db     2954 TTCTCAGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAGGTCAGTTCATTTGCTCCTGAC 3013

Qy      943 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 962
            |||:::|||||||||   ||||||||||:::         ::::::|||||   |||
Db     3014 TATGCCAAAGCCAAAGTATCAGCAGCCCACGTCATCATGATCATTGAAAAAAGCCCTCTG 3073

Qy      963 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 982
            |||||||||   |||         |||   |||||:::|||   ||||||||:::   |||
Db     3074 ATTGACAGCTACAGCCCTCACGGCCTCAAGCCAAATACGTTGGAAGGAAATGTGACATTT 3133

Qy      983 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002
                  ||||||   |||   ||||||   |||||||||:::   :::|||:::||||||||
Db     3134 AATGAGGTCGTGTTCAACTATCCCACTCGACCAGACATCCCCGTGCTCCAGGGGCTGAGC 3193

Qy     1003 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 1022
            |||   :::::::::|||:::|||:::|||   ||||||||||||||||||||||||||
Db     3194 CTCGAGGTGAAGAAGGGCCAGACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGC 3253

Qy     1023 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1042
            |||   ||||||||||||:::|||   |||||||||:::   |||   ||||||   |||
Db     3254 ACAGTTGTTCAGCTCCTAGAGCGCTTCTATGACCCCTTGGCTGGTTCAGTGCTAATTGAT 3313

Qy     1043 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 1062
            |||   :::   |||   |||||||||||||||||||:::   :::   ||||||
Db     3314 GGCAAAGAGATAAAGCACCTGAATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCT 3373

Qy     1063 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082
            |||||||||:::||||||:::|||||||||||||||||||||||||||||||||||||||
Db     3374 CAGGAGCCCATCCTGTTTGACTGCAGCATTGCCGAGAACATTGCCTATGGAGACAACAGC 3433

Qy     1083 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1102
            |||||||||      :::||||||   ||||||||||||   |||
Db     3434 CGGGTCGTATCACATGAAGAGATTATGCAGGCAGCCAAGGAGGCCAACATACACCACTTC 3493

Qy     1103 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122
            ||||||   ||||||||||||||||||:::||||||   ||||||   |||||||||
Db     3494 ATCGAGACACTCCCTGAGAAATACAACACCAGAGTAGGAGACAAAGGAACCCAGCTCTCT 3553

Qy     1123 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 1142
            |||||||||||||||:::||||||||||||||||||:::::::::|||   ||||||
Db     3554 GGTGGCCAGAAACAGCGCATTGCCATAGCTCGCGCTCTTGTTAGACAGCCTCATATTTTG 3613

Qy     1143 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 1162
            ||||||||||||||||||||||||||||       :::||||||||||||||||   |||
Db     3614 CTTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCC 3673

Qy     1163 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 1182
            |||||||||||||||   |||||||||||||:::|||:::   ||||||||||||   |||
Db     3674 CTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATCGCCCACCGCTTGTCCACCATC 3733

Qy     1183 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 1202
            |||||||||||||||||||||||||         |||||||||:::|||||   |||||||||
Db     3734 CAGAATGCAGATTTAATAGTGGTGTTTCAGAATGGCAAAGTCAAGGAGCATGGCACACAT 3793

Qy     1203 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
            |||:::||||||      :::   |||||||||   :::|||:::   |||:::
Db     3794 CAACAGCTGCTGGCCCAGAAAGGCATCTATTTTCCATGGTCAGTGTCCAGGCT 3847
```

RESULT 8
AAD03488
ID AAD03488 standard; cDNA; 4279 BP.
XX
AC AAD03488;
XX
DT 13-JUN-2001 (first entry)
XX
DE Dog P-glycoprotein (PGP) cDNA #1.
XX
KW Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW drug bioavailability; transgenic animal; genetic model; ss.
XX
OS Canis familiaris.
XX
FH Key Location/Qualifiers
FT CDS 17..3862
FT /*tag= a
FT /product= "Dog P-glycoprotein (PGP) #1"
XX
PN WO200123540-A2.
XX
PD 05-APR-2001.
XX
PF 28-SEP-2000; 2000WO-US26767.
XX
PR 28-SEP-1999; 99US-0156510.
XX
PA (GENT-) GENTEST CORP.
XX
PI Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX
DR WPI; 2001-235373/24.
DR P-PSDB; AAE00303.
XX
PT New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT for determining the bioavailability of drugs and for screening for dog
PT PGP inhibitors -
XX

```
PS   Claim 3; Page 58-63; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) cDNA. This
CC   sequence is also referred as Genotype C cDNA. The
CC   PGP enzyme functions as an efflux pump exporting small molecules
CC   across the cell membrane. This enzyme is a member of the ABC
CC   transporter family.
XX
SQ   Sequence 4279 BP; 1294 A; 834 C; 1008 G; 1143 T; 0 other;

Alignment Scores:
Pred. No.:               0          Length:        4279
Score:                   3525.50    Matches:       681
Percent Similarity:      73.05%     Conservative:  238
Best Local Similarity:   54.13%     Mismatches:    270
Query Match:             57.24%     Indels:        69
DB:                      22         Gaps:          9

US-09-873-409-5 (1-1222) x AAD03488 (1-4279)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
Db    227 GTTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy     35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
Db    287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy     45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db    347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy     60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db    407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy     80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db    467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy    100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db    527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy    119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db    587 AAAATTGGAATGTTCTTTCACTCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy    139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db    647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy    159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db    707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy    179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db    767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy    199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
Db    827 GGACAAAAGAAAGAACTTGAA--------------------------------------- 847

Qy    219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
Db    848 ------------------------------------------AGGTACAACAAAAATTTA 865

Qy    239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
Db    866 GAAGAAGCTAAAGGAATTGGGATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCC 925

Qy    259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
Db    926 GCTTTCTTATTGATCTATGCATCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTC 985

Qy    279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
Db    986 CTCTCCAGTGAA------TATTCTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTA 1039

Qy    299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
Db   1040 ATTGGGGCTTTTAGTATTGGACAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGA 1099

Qy    319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
Db   1100 GGAGCAGCTTATGAAATCTTCAAGATAATTGACAATAAACCAAGCATTGACAGCTATTCG 1159

Qy    339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
Db   1160 AAGAGTGGACATAAACCAGATAATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTC 1219

Qy    359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
Db   1220 AGTTACCCTTCTCGAAAAGAAGTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGT 1279

Qy    379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
Db   1280 GGGCAGACAGTGGCGCTGGTTGGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTG 1339

Qy    399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
Db   1340 ATGCAGAGGCTCTATGACCCCACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGG 1399

Qy    419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db   1400 ACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTG 1459

Qy    439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
Db   1460 TTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAG 1519

Qy    459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
Db   1520 ATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTTATCATGAAACTACCTAATAAA 1579

Qy    479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db   1580 TTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATC 1639

Qy    499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db   1640 GCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCA 1699

Qy    519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 538
Db   1700 GCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGC 1759

Qy    539 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558
```

```
Db   1760 CGGACTACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCCGATGTCATTGCT 1819

Qy    559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578
Db   1820 GGTTTTGATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAG 1879

Qy    579 GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 598
Db   1880 GGCATTTACTTCAAACTTGTCACAATGCAG---ACAAGAGGAAATGAAATTGAGTTAGAA 1936

Qy    599 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 618
Db   1937 AATGCCACTGGTGAATCCAAAAGTGAAAGTGATGCCTTGGAAATG---TCTCCAAAAGAT 1993

Qy    619 IleLysSerAspPheIle------------------------------------------ 624
Db   1994 TCAGGGTCCAGTTTAATAAAAAGAAGATCAACTCGCAGGAGTATACATGCACCACAAGGC 2053

Qy    625 ---AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643
Db   2054 CAAGACAGAAAGCTTGGTACAAAAGAGGACTTGAATGAGAATGTACCTCCAGTTTCCTTC 2113

Qy    644 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 663
Db   2114 TGGAGGATTCTGAAGCTGAACTCAACTGAATGGCCTTATTTTGTGGTTGGTATATTTTGT 2173

Qy    664 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 683
Db   2174 GCTATTATAAACGGAGGCCTGCAACCAGCATTTTCAATAATATTTTCAAGGATTATAGGG 2233

Qy    684 MetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 702
Db   2234 ATCTTTACCCGAGATGAGGATCCTGAAACAAAACGACAGAATAGTAACATGTTTTCTGTA 2293

Qy    703 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722
Db   2294 TTGTTTCTAGTCCTTGGAATTATTTCTTTTATTACATTTTTCCTCCAGGGCTTCACATTT 2353

Qy    723 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 742
Db   2354 GGCAAAGCTGGGGAGATCCTCACTAAGCGGCTTCGATACATGGTTTTCAGATCCATGCTG 2413

Qy    743 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIle 762
Db   2414 AGACAGGATGTCAGCTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACAACCAGG 2473

Qy    763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 782
Db   2474 CTTGCCAATGATGCGGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTCATTACC 2533

Qy    783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 802
Db   2534 CAGAATATAGCAAATCTTGGGACAGGCATTATTATATCCTTAATCTATGGTTGGCAATTA 2593

Qy    803 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 822
Db   2594 ACACTTTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAA 2653

Qy    823 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 842
Db   2654 ATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCT 2713

Qy    843 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 862
Db   2714 ACAGAAGCCATCGAAAACTTCCGAACTGTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAA 2773

Qy    863 GlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 882
Db   2774 TACATGTATGCACAGAGTTTGCAAGTACCATACAGAAACTCTTTGAGGAAAGCACACATC 2833

Qy    883 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 902
Db   2834 TTCGGGGTCTCATTTTCTATCACCCAGGCAATGATGTATTTTTCCTATGCTGGCTGTTTC 2893

Qy    903 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 922
Db   2894 CGGTTTGGTGCCTACTTGGTGGCAAATGAGTTCATGAACTTTCAGGATGTTCTTTTGGTA 2953

Qy    923 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 942
Db   2954 TTCTCAGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAGGTCAGTTCATTTGCTCCTGAC 3013

Qy    943 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 962
Db   3014 TATGCCAAAGCCAAAGTATCAGCAGCCCACGTCATCATGATCATTGAAAAAAGCCCTCTG 3073

Qy    963 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 982
Db   3074 ATTGACAGCTACAGCCCTCACGGCCTCAAGCCAAATACGTTGGAAGGAAATGTGACATTT 3133

Qy    983 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002
Db   3134 AATGAGGTCGTGTTCAACTATCCCACTCGACCAGACATCCCCGTGCTCCAGGGGCTGAGC 3193

Qy   1003 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 1022
Db   3194 CTCGAGGTGAAGAAGGGCCAGACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGC 3253

Qy   1023 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1042
Db   3254 ACAGTTGTTCAGCTCCTAGAGCGCTTCTATGACCCCTTGGCTGGTTCAGTGCTAATTGAT 3313

Qy   1043 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 1062
Db   3314 GGCAAAGAGATAAAGCACCTGAATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCT 3373

Qy   1063 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082
Db   3374 CAGGAGCCCATCCTGTTTGACTGCAGCATTGCCGAGAACATTGCCTATGGAGACAACAGC 3433

Qy   1083 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1102
Db   3434 CGGGTCGTATCACATGAAGAGATTATGCAGGCAGCCAAGGAGGCCAACATACACCACTTC 3493

Qy   1103 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122
Db   3494 ATCGAGACACTCCCTGAGAAATACAACACCAGAGTAGGAGACAAAGGAACCCAGCTCTCT 3553

Qy   1123 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 1142
Db   3554 GGTGGCCAGAAACAGCGCATTGCCATAGCTCGCGCTCTTGTTAGACAGCCTCATATTTTG 3613

Qy   1143 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 1162
Db   3614 CTTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCC 3673

Qy   1163 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 1182
Db   3674 CTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATCGCCCACCGCTTGTCCACCATC 3733

Qy   1183 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 1202
Db   3734 CAGAATGCAGATTTAATAGTGGTGTTTCAGAATGGCAAAGTCAAGGAGCATGGCACACAT 3793

Qy   1203 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db   3794 CAACAGCTGCTGGCCCAGAAAGGCATCTATTTTTCCATGGTCAGTGTCCAGGCT 3847
```

```
RESULT 9
AAD03489
ID   AAD03489 standard; cDNA; 4317 BP.
XX
AC   AAD03489;
XX
DT   13-JUN-2001  (first entry)
```

```
XX
DE   Dog P-glycoprotein (PGP) cDNA #2.
XX
KW   Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             70..3912
FT                   /*tag=  a
FT                   /product= "Dog P-glycoprotein (PGP) #2"
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00304.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 1; Page 66-72; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) cDNA. The
CC   PGP enzyme functions as an efflux pump exporting small molecules
CC   across the cell membrane. This enzyme is a member of the ABC
CC   transporter family.
XX
SQ   Sequence 4317 BP; 1293 A; 844 C; 1019 G; 1161 T; 0 other;

Alignment Scores:
Pred. No.:               0          Length:         4317
Score:                   3525.50    Matches:        680
Percent Similarity:      73.05%     Conservative:   239
Best Local Similarity:   54.05%     Mismatches:     270
Query Match:             57.24%     Indels:         69
DB:                      22         Gaps:           9

US-09-873-409-5 (1-1222) x AAD03489 (1-4317)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
              |||:::::|||   :::|||::::::::::::||||||   |||||||||||   |||
Db      217 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 276

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
              |||   |||   |||:::|||:::       :::      :::
Db      277 GTTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 336

Qy       35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                                  :::          ||||||      |||
Db      337 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 396

Qy       45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                               ||||||   |||||||||            |||:::
Db      397 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 456

Qy       60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
               |||||||||:::|||   |||   :::   |||   ||||||         :::||||||
Db      457 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 516

Qy       80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
              |||||||||||||||:::::::::   |||:::||||||||||||||||||    |||:::
Db      517 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 576

Qy      100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
              ||||||||||||||||||:::|||   |||:::   ||||||::::::||||||||||||
Db      577 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 636

Qy      119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
              |||:::   :::   ||||||:::::::::||||||         |||      ||||||
Db      637 AAAGTTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 696

Qy      139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
                 |||||||||||||||||||||   |||:::   ||||||::::::         |||
Db      697 ACACCTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 756

Qy      159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
              ||||||      ::::::::::::   |||   |||   |||||||||   ||||||:::
Db      757 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 816

Qy      179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
              |||||||||||||||||||||||||||::::::||||||||||||||||||||||||
Db      817 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 876

Qy      199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
                 |||:::|||||||||:::
Db      877 GGACAAAAGAAAGAACTTGAA--------------------------------------- 897

Qy      219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
                                                    ||||||   :::||||||
Db      898 ------------------------------------------AGGTACAACAAAAATTTA 915

Qy      239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
              ::::::||||||      |||||||||:::   |||   :::   :::|||:::||||||
Db      916 GAAGAAGCTAAAAGAATTGGGATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCC 975

Qy      259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
                 :::      :::      :::|||   |||||||||||||||||||||||||:::
Db      976 GCTTTCTTATTGATCTATGCATCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTC 1035

Qy      279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
              |||:::   |||      |||||||||||   ||||||   ||||||||||||||||:::
Db     1036 CTCTCCAGTGAA------TATACTATTGGACAGGTACTCACTGTCTTCTTTTCTGTATTA 1089

Qy      299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
                 ::::::   ||||||   |||   |||         |||   ||||||   ||||||
Db     1090 ATTGGGGCTTTTAGTATTGGACAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGA 1149

Qy      319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
              |||||||||:::      ||||||:::::||||||   |||||||||||||||::::::|||
Db     1150 GGAGCAGCTTATGAAATCTTCAAGATAATTGACAATAAACCAAGCATTGACAGCTATTCG 1209

Qy      339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
                 :::|||:::||||||::::::|||:::|||   :::||||||||||||||   |||
Db     1210 AAGAGTGGACATAAACCAGATAATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTC 1269

Qy      359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
              :::||||||||||||      :::|||||||||||||||||||||||||:::::::::|||
Db     1270 AGTTACCCTTCTCGAAAAGAAGTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGT 1329

Qy      379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
              |||:::||||||||||||||||||   :::|||   ||||||||||||   ||||||||
```

```
Db   1330 GGGCAGACAGTGGCGCTGGTTGGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTG 1389

Qy    399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
Db   1390 ATGCAGAGGCTCTATGACCCCACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGG 1449

Qy    419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db   1450 ACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTG 1509

Qy    439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
Db   1510 TTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAG 1569

Qy    459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
Db   1570 ATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTTATCATGAAACTACCTAATAAA 1629

Qy    479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db   1630 TTTGACACTCTGGTTGGAGAGAGAGGGGCCCGGCTGAGTGGTGGACAGAAACAGAGAATC 1689

Qy    499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db   1690 GCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCA 1749

Qy    519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 538
Db   1750 GCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGC 1809

Qy    539 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558
Db   1810 CGGACTACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCCGATGTCATTGCT 1869

Qy    559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578
Db   1870 GGTTTTGATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAG 1929

Qy    579 GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 598
Db   1930 GGCATTTACTTCAAACTTGTCACAATGCAG---ACAAGAGGAAATGAAATTGAGTTAGAA 1986

Qy    599 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 618
Db   1987 AATGCCACTGGTGAATCCAAAAGTGAAAGTGATGCCTTGGAAATG---TCTCCAAAAGAT 2043

Qy    619 IleLysSerAspPheIle------------------------------------------ 624
Db   2044 TCAGGGTCCAGTTTAATAAAAAGAAGATCAACTCGCAGGAGTATACATGCACCACAAGGC 2103

Qy    625 ---AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643
Db   2104 CAAGACAGAAAGCTTGGTACAAAAGAGGACTTGAATGAGAATGTTCCTTCAGTTTCCTTC 2163

Qy    644 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 663
Db   2164 TGGAGGATTCTGAAGCTGAACTCAACTGAATGGCCTTATTTTGTGGTTGGTATATTTTGT 2223

Qy    664 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 683
Db   2224 GCTATTATAAACGGAGGCCTGCAACCAGCATTTTCAATAATATTTTCAAGGATTATAGGG 2283

Qy    684 MetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 702
Db   2284 ATCTTTACCCGAGATGAGGATCCTGAAACAAAACGACAGAATAGTAACATGTTTTCTGTA 2343

Qy    703 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722
Db   2344 TTGTTTCTAGTCCTTGGAATTATTTCTTTTATTACATTTTTCCTCCAGGGCTTCACATTT 2403

Qy    723 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 742
Db   2404 GGCAAAGCTGGGGAGATCCTCACTAAGCGGCTTCGATACATGGTTTTCAGATCCATGCTG 2463

Qy    743 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIle 762
Db   2464 AGACAGGATGTCAGCTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACAACCAGG 2523

Qy    763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 782
Db   2524 CTTGCCAATGATGCGGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTCATTACC 2583

Qy    783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 802
Db   2584 CAGAATATAGCAAATCTTGGGACAGGCATTATTATATCCTTAATCTATGGTTGGCAATTA 2643

Qy    803 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 822
Db   2644 ACACTTTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAA 2703

Qy    823 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 842
Db   2704 ATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCT 2763

Qy    843 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 862
Db   2764 ACAGAAGCCATCGAAAACTTCCGAACTGTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAA 2823

Qy    863 GlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 882
Db   2824 TACATGTATGCACAGAGTTTGCAAGTACCATACAGAAACTCTTTGAGGAAAGCACACATC 2883

Qy    883 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 902
Db   2884 TTCGGGGTCTCATTTTCTATCACCCAGGCAATGATGTATTTTTCCTATGCTGGCTGTTTC 2943

Qy    903 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 922
Db   2944 CGGTTTGGTGCCTACTTGGTGGCAAATGAGTTCATGAACTTTCAGGATGTTCTTTTGGTA 3003

Qy    923 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 942
Db   3004 TTCTCAGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAGGTCAGTTCATTTGCTCCTGAC 3063

Qy    943 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 962
Db   3064 TATGCCAAAGCCAAAGTATCAGCAGCCCACGTCATCATGATCATTGAAAAAAGCCCTCTG 3123

Qy    963 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 982
Db   3124 ATTGACAGCTACAGCCCTCACGGCCTCAAGCCAAATACGTTGGAAGGAAATGTGACATTT 3183

Qy    983 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002
Db   3184 AATGAGGTCGTGTTCAACTATCCCACTCGACCAGACATCCCCGTGCTCCAGGGGCTGAGC 3243

Qy   1003 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 1022
Db   3244 CTCGAGGTGAAGAAGGGCCAGACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGC 3303

Qy   1023 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1042
Db   3304 ACAGTTGTTCAGCTCCTAGAGCGCTTCTATGACCCCTTGGCTGGTTCAGTGCTAATTGAT 3363

Qy   1043 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 1062
Db   3364 GGCAAAGAGATAAAGCACCTGAATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCT 3423

Qy   1063 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082
Db   3424 CAGGAGCCCATCCTGTTTGACTGCAGCATTGCCGAGAACATTGCCTATGGAGACAACAGC 3483

Qy   1083 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1102
Db   3484 CGGGTCGTATCACATGAAGAGATTATGCAGGCAGCCAAGGAGGCCAACATACACCACTTC 3543
```

```
Qy   1103 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122
          ||||||   ||||||||||||||||||||||:::||||||   ||||||  |||||||||
Db   3544 ATCGAGACACTCCCTGAGAAATACAACACCAGAGTAGGAGACAAAGGAACCCAGCTCTCT 3603

Qy   1123 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 1142
          ||||||||||||||||||:::||||||||||||||||||:::::::::|||  ||||||
Db   3604 GGTGGCCAGAAACAGCGCATTGCCATAGCTCGCGCTCTTGTTAGACAGCCTCATATTTTG 3663

Qy   1143 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 1162
          |||||||||||||||||||||||||||||   :::||||||||||||||||||   |||
Db   3664 CTTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCC 3723

Qy   1163 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 1182
          |||||||||||||||   |||||||||||:::|||:::  ||||||||||||   |||
Db   3724 CTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATCGCCCACCGCTTGTCCACCATC 3783

Qy   1183 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 1202
          ||||||||||||||||||||||||      |||||||||:::||||||   |||||||||
Db   3784 CAGAATGCAGATTTAATAGTGGTGTTTCAGAATGGCAAAGTCAAGGAGCATGGCACACAT 3843

Qy   1203 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
          |||:::||||||      :::   |||||||||   :::::::::   |||:::
Db   3844 CAACAGCTGCTGGCTCAGAAAGGCATCTATTTTTCCATGATCAGTGTCCAGGCT 3897
```

RESULT 10
AAF86127
ID   AAF86127 standard; cDNA; 4186 BP.
XX
AC   AAF86127;
XX
DT   25-JUN-2001 (first entry)
XX
DE   Cynomologous monkey P-glycoprotein cDNA.
XX
KW   Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   efflux pump; ss.
XX
OS   Macaca fascicularis.
XX
FH   Key   Location/Qualifiers
FT   CDS   100..3942
FT   /*tag= a
FT   /product= "PGP"
FT   /note= "P-glycoprotein"
XX
PN   WO200123565-A1.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26592.
XX
PR   28-SEP-1999;   99US-0156921.
PR   12-OCT-1999;   99US-0158818.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ, Steimel-Crespi DT, Crespi CL;
XX
DR   WPI; 2001-316136/33.
DR   P-PSDB; AAB81064.
XX
PT   Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein
PT   (PGP) and homologous PGP polypeptides are useful for predicting
PT   bioavailability of compound and increasing PGP transporter activity in
PT   cell -
XX
PS   Example 1; Page 51-57; 84pp; English.
XX
CC   This invention relates to a polynucleotide sequence encoding a
CC   cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the
CC   PGP protein. PGP, also known as multidrug transporter, MDR1 is a member
CC   of the ABC transporter superfamily. The enzyme serves as an efflux pump
CC   exporting small molecules across the cell membrane. The invention
CC   includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence
CC   and protein, and also that of an allelic variant. The PGP polynucleotide
CC   sequence is useful for increasing PGP transporter activity in a cell.
CC   Antisense sequences of the cDNA are useful for inhibiting PGP transport
CC   activity in a mammalian cell. They may also be used for increasing the
CC   bioavailability of a drug. The present sequence represents cDNA encoding
CC   cynomologous monkey P-glycoprotein.
XX
SQ   Sequence 4186 BP; 1226 A; 799 C; 1039 G; 1122 T; 0 other;

Alignment Scores:			
Pred. No.:	0	Length:	4186
Score:	3524.50	Matches:	679
Percent Similarity:	73.75%	Conservative:	248
Best Local Similarity:	54.02%	Mismatches:	261
Query Match:	57.23%	Indels:	69
DB:	22	Gaps:	11

US-09-873-409-5 (1-1222) x AAF86127 (1-4186)

```
Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          |||::::::|||   ||||||::::::::::||||||   ||||||||||||   |||
Db    250 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGAGCTGGACTTCCTCTCATGATGCTG 309

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSer---------------------GlyCys 33
          |||   |||:::|||:::|||           :::                     |||
Db    310 GTGTTTGGAGACATGACGGATACCTTTGCAAATGCAGGAAATTTAGGAGATTTAGGAGCT 369

Qy     34 LeuValGlnThrAsnThrTyrSer------------------------------------ 41
          |||:::   ||||||:::   :::
Db    370 CTGTTG---ACTAATAGCAGTAATATCACTGATACAGTGCCCGTCATGAATCTGGAGGAA 426

Qy     42 ---PhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
             |||         ||||||   |||||||||         |||:::
Db    427 GATATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCT 486

Qy     61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
          |||||||||:::|||   |||   :::   ||||||         :::|||||||||
Db    487 TACATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAA 546

Qy     81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
          ||||||||||||:::::::::   |||:::|||||||||||||||      |||:::|||
Db    547 CAGTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGG 606

Qy    101 GluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLys 119
          |||||||||||||||:::|||   |||:::   ||||||::::::||||||||||||||
Db    607 GAGCTTAACACCCGGCTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAA 666

Qy    120 IleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuVal 139
          |||   :::   ||||||:::|||:::||||||         |||      ||||||
Db    667 ATTGGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACA 726

Qy    140 LysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAla 159
          :::||||||||||||||||||||||||   |||:::   ||||||::::::      ||||||
Db    727 CGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCT 786

Qy    160 AlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLys 179
          |||      ::::::::::::   |||   |||   |||||||||   ||||||:::|||
Db    787 GCAGTCTGGGCAAAGATACTGTCTTCATTTACTGATAAAGAACTCTTAGCTTATGCAAAA 846

Qy    180 AlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAla 199
          ||||||||||||||||||||||||||||::::::||||||||||||||||||||||
Db    847 GCTGGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGA 906

Qy    200 GlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyr 219
          |||:::|||||||||:::
Db    907 CAAAAGAAAGAACTCGAA------------------------------------------ 924

Qy    220 PheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLys 239
                                                  ||||||   :::|||||||:::
Db    925 ----------------------------------------AGGTACAACAAAAATTTAGAA 945

Qy    240 AspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaVal 259
```

```
Db    946 GAAGCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCT 1005
Qy    260 TyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeu 279
Db   1006 TTCCTGCTTATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTC 1065
Qy    280 AsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHis 299
Db   1066 TCAAAGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATT 1119
Qy    300 SerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGly 319
Db   1120 GGGGCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGA 1179
Qy    320 AlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThr 339
Db   1180 GCAGCTTTTGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAG 1239
Qy    340 AlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsn 359
Db   1240 AGTGGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGT 1299
Qy    360 TyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGly 379
Db   1300 TACCCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGG 1359
Qy    380 GluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeu 399
Db   1360 CAGACGGTGGCCCTGGTTGGAAACAGCGGCTGTGGGAAGAGCACAACGGTCCAGCTGATG 1419
Qy    400 GlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAla 419
Db   1420 CAGAGGCTTTATGACCCCACAGAGGGCATGGTCAGTGTTGATGGACAGGATATTAGGACC 1479
Qy    420 LeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPhe 439
Db   1480 ATAAACGTAAGGTTTCTACGGGAAATCATCGGTGTGGTGAGTCAGGAACCTGTATTGTTT 1539
Qy    440 GlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMet 459
Db   1540 GCCACCACGATAGCTGAAAACATTCGCTATGGTCGTGAAGATGTCACCATGGATGAGATT 1599
Qy    460 GluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPhe 479
Db   1600 GAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCAGAAATTT 1659
Qy    480 AsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAla 499
Db   1660 GACACCCTGGTTGGAGAGAGAGGGGCCCAGCTGAGTGGTGGGCAGAAGCAGAGGATCGCC 1719
Qy    500 IleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAla 519
Db   1720 ATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGACGAGGCCACGTCAGCC 1779
Qy    520 LeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArg 539
Db   1780 TTGGACACAGAAAGTGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGG 1839
Qy    540 ThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThr 559
Db   1840 ACCACCATTGTGATAGCTCATCGTTTGTCTACGGTTCGTAATGCCGACGTCATCGCTGGT 1899
Qy    560 LeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGly 579
Db   1900 TTCGATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAGCTCATGAAAGAGAAAGGC 1959
Qy    580 LeuTyrTyrSerLeuValMetSerGlnAspIle--------------------LysLys 592
Db   1960 ATTTACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAATTGAATTAGAAAATGCA 2019
Qy    593 AlaAspGlu---------GlnMetGluSerMetThrTyrSerThrGluArgLysThrAsn 609
Db   2020 GCTGATGAATCCAAAAGTGAAATTGATACCTTGGAAATGTCTTCACATGATTCAGGATCC 2079
Qy    610 SerLeu---------------ProLeuHisSerValLysSerIleLysSerAspPheIle 624
Db   2080 AGTCTAATAAGAAAAAGATCCACTCGTAGGAGTGTCCGTGGATCACAAGGCCAA------ 2133
Qy    625 AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 644
Db   2134 GACAGAAAGCTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGG 2193
Qy    645 LysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 664
Db   2194 AGGATTATGAAGCTAAATTTAACTGAGTGGCCTTATTTTGTTGTTGGTGTATTTTGTGCC 2253
Qy    665 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 684
Db   2254 ATTATAAATGGAGGTCTGCAACCAGCATTTGCAGTAATATTTTCAAAGATTATAGGGATT 2313
Qy    685 Phe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 703
Db   2314 TTTACAAGAAATGATGATGCCGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTG 2373
Qy    704 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 723
Db   2374 TTTCTAGTCCTTGGAATTGTTTCTTTTATTACATTTTTCCTTCAGGGCTTCACATTTGGC 2433
Qy    724 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 743
Db   2434 AAAGCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGA 2493
Qy    744 GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 763
Db   2494 CAGGATGTGAGCTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTC 2553
Qy    764 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 783
Db   2554 GCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTATAATTACCCAG 2613
Qy    784 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr 803
Db   2614 AATATAGCAAATCTTGGGACAGGAATAATTATATCCTTAATCTATGGTTGGCAACTGACA 2673
Qy    804 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 823
Db   2674 CTGTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATG 2733
Qy    824 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 843
Db   2734 TTGTCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACT 2793
Qy    844 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 863
Db   2794 GAAGCAATAGAAAACTTCCGAACTGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACAT 2853
Qy    864 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 883
Db   2854 ATGTATGATCAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTT 2913
Qy    884 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 903
Db   2914 GGAATCACGTTTTCCTTCACGCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGG 2973
Qy    904 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 923
Db   2974 TTTGGAGCCTACTTGGTGGCACATAGTCTCATGAGCTTTGAGGATGTTCTGTTAGTATTT 3033
Qy    924 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 943
Db   3034 TCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTAT 3093
Qy    944 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 963
Db   3094 GCCAAAGCCAAAGTATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATT 3153
```

```
Qy     964 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 983
           ||||||   |||   |||||   ||||||:::|||  |||||||||:::    |||
Db    3154 GACAGCTACAGCACAGAAGGCCTAAAGCCGAACACATTGGAAGGAAATGTCACATTTAAT 3213

Qy     984 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 1003
           ||||||   |||   |||||   |||   |||:::   :::|||:::||||||||||||
Db    3214 GAAGTTGTATTCAACTATCCCACCCGACTGGACATCCCAGTGCTTCAGGGGCTGAGCCTG 3273

Qy    1004 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 1023
              :::::::::|||:::|||:::|||   ||||||||||||||||||||||||||||
Db    3274 GAAGTGAAGAAGGGCCAGACGCTGGCCCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACG 3333

Qy    1024 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 1043
              ||||||||||||:::|||   |||||||||:::   |||:::||||||   ||||||
Db    3334 GTGGTCCAGCTCCTGGAGCGGTTCTATGACCCCTTGGCGGGGAAAGTGCTGCTTGACGGC 3393

Qy    1044 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 1063
              :::   |||:::||||||||||||||||||||||||:::   :::   ||||||   |||
Db    3394 AAAGAAATAAAGCAACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAG 3453

Qy    1064 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 1083
           ||||||:::||||||:::|||||||||:::||||||||||||||||||||||||||||
Db    3454 GAGCCCATCCTGTTTGACTGCAGCATTAGTGAGAACATTGCCTATGGAGACAACAGCCGG 3513

Qy    1084 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 1103
           ||||||      :::||||||      ||||||      ||||||||||||:::||||||
Db    3514 GTGGTGTCACAGGAAGAGATCGTGAGGGCAGCCAAGGAGGCCAATATACACGCCTTCATC 3573

Qy    1104 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 1123
           |||   ||||||   ||||||:::|||:::||||||   ||||||   ||||||||||||
Db    3574 GAGTCACTGCCTAATAAATATAGCACCAGAGTAGGAGACAAAGGAACTCAGCTCTCTGGT 3633

Qy    1124 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 1143
           ||||||||||||||||:::|||||||||||||||||||:::::::::|||   ||||||||
Db    3634 GGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTT 3693

Qy    1144 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 1163
           ||||||||||||||||||||||||||||   :::||||||||||||||||||   ||||||
Db    3694 TTGGATGAAGCCACATCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTG 3753

Qy    1164 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 1183
           ||||||||||||   ||||||||||||:::|||:::   ||||||||||||   ||||||
Db    3754 GACAAAGCCAGAGAAGGCCGTACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAG 3813

Qy    1184 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 1203
           ||||||||||||||||||||||      ||||||::::::||||||   ||||||||||||
Db    3814 AATGCAGACTTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCACGGCACACATCAG 3873

Qy    1204 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
           :::||||||      :::   |||||||||   :::|||:::   |||:::
Db    3874 CAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3924
```

RESULT 11
AAZ49333
ID   AAZ49333 standard; cDNA; 3860 BP.
XX
AC   AAZ49333;
XX
DT   14-MAR-2000 (first entry)
XX
DE   Human G185V mutant multidrug resistance-1 (MDR-1) cDNA.
XX
KW   Multidrug resistance; MDR-1; P-glycoprotein;
KW   transmembrane efflux pump; haematopoietic stem cell; transduction;
KW   bone marrow transplantation; chemotherapy; radiation therapy; cancer;
KW   gene therapy; gene replacement; genetic defect; thalassaemia;
KW   Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
KW   cytokine; mutant; ds.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..3843
FT   /*tag= a
FT   /product= "Human G185V mutant MDR-1 protein"
XX
PN   WO9961589-A2.
XX
PD   02-DEC-1999.
XX
PF   27-MAY-1999;   99WO-US11825.
XX
PR   28-MAY-1998;   98US-0086988.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI   Sorrentino B, Bunting K;
XX
DR   WPI; 2000-072615/06.
DR   P-PSDB; AAY58187.
XX
PT   Ex vivo expansion of hematopoietic stem cells transduced with a
PT   sequence encoding human multidrug resistance-1, used for bone marrow
PT   transplantation -
XX
PS   Example 1; Page 79-82; 113pp; English.
XX
CC   This sequence represents cDNA encoding human G185V mutant multidrug
CC   resistance protein MDR-1, where the the Gly residue at postion 185
CC   of the wild-type protein (AAY58186) is replaced by Val. MDR-1 is a
CC   transmembrane efflux pump, responsible for the export of drugs from
CC   cells, particularly cancer cells. The wild-type MDR-1 shows increased
CC   resistance to etoposide and decreased resistance to vinca alkaloids
CC   compared with the G185V mutant. The invention relates to transducing
CC   haematopoietic stem cells with nucleic acid encoding an MDR protein
CC   and culturing the modified cells. The modified haematopoietic stem
CC   cells are useful in bone marrow transplantation (to reconstitute
CC   haematopoietic systems in patients who have undergone chemotherapy or
CC   radiation therapy) and in ex vivo gene therapy of genetic defects in
CC   cells derived from haematopoietic stem cells, e.g., thalassaemia,
CC   Gauchier's disease, sickle cell anaemia or leukaemia. The modified
CC   cells can also be used to identify factors involved in regulating
CC   proliferation and differentiation in haematopoietic stem cells.
CC   Haematopoietic stem cells that express MDR-1 will be protected against
CC   chemotherapeutic agents, so can be engrafted while the patient is
CC   undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells
CC   provides sufficient cells to permit standard biochemical analysis.
CC   Overexpression of MDR-1 allows cytokine-driven expansion of
CC   haematopoietic stem cells by at least 10-fold compared with a maximum
CC   of 4-fold in known procedures.
XX
SQ   Sequence 3860 BP; 1135 A; 746 C; 956 G; 1023 T; 0 other;

Alignment Scores:			
Pred. No.:	0	Length:	3860
Score:	3521.50	Matches:	681
Percent Similarity:	73.92%	Conservative:	246
Best Local Similarity:	54.31%	Mismatches:	264
Query Match:	57.18%	Indels:	63
DB:	21	Gaps:	10

US-09-873-409-5 (1-1222) x AAZ49333 (1-3860)

```
Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||::::::|||   ||||||::::::::::::||||||   ||||||||||||   |||
Db     151 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210

Qy      21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
           |||   |||||||||:::|||                           :::|||:::|||
Db     211 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 270

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                    ::::::      |||         ||||||
Db     271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330
```

```
Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    331 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 390

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    451 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 510

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    511 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 570

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 630

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    631 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    691 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db    811 AAGAAAGAACTTGAA--------------------------------------------- 825

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db    826 ------------------------------------AGGTACAACAAAAATTTAGAAGAA 849

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db    850 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 909

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db    910 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 969

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db    970 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1023

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db   1024 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1083

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   1084 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1143

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   1144 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1203

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   1204 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1263

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   1264 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1323

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   1324 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1383

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   1384 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1443

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   1444 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1503

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   1504 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1563

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   1564 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 1623

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   1624 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 1683

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db   1684 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACC 1743

Qy    541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db   1744 ACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC 1803

Qy    561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db   1804 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATT 1863

Qy    581 TyrTyrSerLeuValMetSerGln------------------AspIleLysLysAlaAsp 594
Db   1864 TACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 1923

Qy    595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
Db   1924 GATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGT 1983

Qy    615 SerVal------------------LysSerIleLys---SerAspPheIleAspLysAla 627
Db   1984 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAG 2043

Qy    628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
Db   2044 CTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATG 2103

Qy    648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
Db   2104 AAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAAT 2163

Qy    668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
Db   2164 GGAGGCCTGCAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGA 2223

Qy    688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db   2224 ATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCC 2283

Qy    707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db   2284 CTTGGAATTATTTCTTTTATTACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGA 2343

Qy    727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db   2344 GAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTG 2403

Qy    747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
```

```
                :::|||||||||:::    :::|||:::||||||  |||||||||  ||||||   |||
Db    2404 AGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGAT 2463

Qy     767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
               |||||::::::||||||  |||||||||:::   |||:::|||||||||
Db    2464 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCA 2523

Qy     787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
           |||:::|||        :::|||||||||||||||||||||||||::::::|||   |||:::
Db    2524 AATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTC 2583

Qy     807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
           |||:::|||   |||::::::|||:::   |||:::::::|||         ::::::|||
Db    2584 TTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGA 2643

Qy     827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
              |||   ||||||||:::||||||:::   ||||||||||||||||||||||||:::
Db    2644 CAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 2703

Qy     847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
           ||||||   ||||||:::||||||||||||||:::|||:::      ||||||   ||||||
Db    2704 GAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCT 2763

Qy     867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
           :::   ||||||      :::||||||:::   :::||||||   |||   |||
Db    2764 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 2823

Qy     887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
           ::::::|||:::   |||   :::||||||:::||||||         |||||||||||||||
Db    2824 TTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCC 2883

Qy     907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
           ||||||:::            |||:::   |||   :::   :::||||||:::|||:::
Db    2884 TACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTT 2943

Qy     927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
              :::|||||||||||||:::|||:::               ||||||:::|||:::||||||
Db    2944 GTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCC 3003

Qy     947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
           |||      |||||||||:::         ::::::||||||   |||   |||||||||
Db    3004 AAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTAC 3063

Qy     967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
           |||   ||||||         |||:::|||   |||||||||:::   |||   ||||||
Db    3064 AGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTA 3123

Qy     987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
           |||   ||||||   |||||||||:::   :::|||:::|||||||||||||   ::::::
Db    3124 TTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAG 3183

Qy    1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
           :::|||:::|||:::|||   ||||||||||||||||||||||||||||||   ||||||
Db    3184 AAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAG 3243

Qy    1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
           ||||||:::|||   |||||||||:::   |||:::||||||   ||||||      :::
Db    3244 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATA 3303

Qy    1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
           |||   ||||||||||||||||||||:::   :::   ||||||   |||||||||:::
Db    3304 AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATC 3363

Qy    1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
           ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3364 CTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCA 3423

Qy    1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
              :::||||||      ||||||      ||||||||||||:::||||||||   |||
Db    3424 CAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTG 3483

Qy    1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
           |||   ||||||:::|||:::||||||   ||||||   ||||||||||||||||||||
Db    3484 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAA 3543

Qy    1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
           ||||||:::||||||||||||||||||||:::::::::|||   ||||||||||||||||||
Db    3544 CAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAA 3603

Qy    1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
           ||||||||||||||||||   :::||||||||||||||||||   |||||||||||||||
Db    3604 GCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCC 3663

Qy    1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
           |||   ||||||||||||:::|||:::      ||||||||||||   ||||||||||||||||
Db    3664 AGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGAC 3723

Qy    1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
           ||||||||||||      ||||||::::::||||||   ||||||||||||:::||||||
Db    3724 TTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTG 3783

Qy    1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
              :::   |||||||||   :::|||:::   |||:::
Db    3784 GCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3825
```

RESULT 12
ABA94366
ID   ABA94366 standard; DNA; 3860 BP.
XX
AC   ABA94366;
XX
DT   26-MAR-2002 (first entry)
XX
DE   Human BCRP DNA related seq Id No. 3.
XX
KW   Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;
KW   hepatotropic; neurodegenerative; cytostatic; antianemic; muscular; BCRP;
KW   cardiant; gene therapy; ds.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FT   CDS   1..3843
FT   /*tag= a
XX
PN   WO200192877-A2.
XX
PD   06-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US17459.
XX
PR   31-MAY-2000; 2000US-0584586.
PR   29-MAY-2001; 2001US-0866866.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI   Sorrentino B, Schuetz J;
XX
DR   WPI; 2002-114368/15.
DR   P-PSDB; ABB07267.
XX
PT   Identifying a stem cell, for treating e.g., muscular dystrophy,
PT   myocardial infarction, Parkinson's disease, or neurodegenerative
PT   disorders, comprises detecting the expression of an ATP transport
PT   protein (BCRP) by a cell -
XX
PS   Disclosure; Page 59-60; 87pp; English.
XX
CC   The invention provides a method of identifying and/or isolating a stem
CC   cell that involves detecting the expression of an ATP transport protein
CC   containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
CC   comprising stem cells. The isolated stem cells may be used in the
CC   treatment of diseases such as muscular dystrophy, degenerative liver
CC   disorder, myocardial infarction, Parkinson's disease, degenerative
CC   disorders of the brain, and for tissue regeneration or replacement.
CC   Haematopoietic cells can be used in bone marrow transplants (e.g., for

```
CC   treatment of leukemia) and for ex vivo gene therapy for treating blood
CC   diseases such as sickle cell anemia and thalassemia. The stem cells can
CC   also be used as cell targets in gene therapy protocols. The present
CC   sequence represents a sequence related to the BCRP for which no relevant
CC   information has been provided in the specification.
XX
SQ   Sequence 3860 BP; 1135 A; 746 C; 956 G; 1023 T; 0 other;

Alignment Scores:
Pred. No.:              0           Length:         3860
Score:                  3521.50     Matches:        681
Percent Similarity:     73.92%      Conservative:   246
Best Local Similarity:  54.31%      Mismatches:     264
Query Match:            57.18%      Indels:         63
DB:                     24          Gaps:           10

US-09-873-409-5 (1-1222) x ABA94366 (1-3860)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    151 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210

Qy     21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db    211 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 270

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db    271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    331 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 390

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    451 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 510

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    511 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 570

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 630

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    631 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    691 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db    811 AAGAAAGAACTTGAA--------------------------------------------- 825

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db    826 -------------------------------------AGGTACAACAAAAATTTAGAAGAA 849

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db    850 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 909

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db    910 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 969

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db    970 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1023

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db   1024 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1083

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   1084 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1143

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   1144 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1203

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   1204 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1263

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   1264 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1323

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   1324 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1383

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   1384 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1443

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   1444 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1503

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   1504 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1563

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   1564 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 1623

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   1624 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 1683

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db   1684 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACC 1743

Qy    541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db   1744 ACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC 1803

Qy    561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db   1804 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATT 1863

Qy    581 TyrTyrSerLeuValMetSerGln------------------AspIleLysLysAlaAsp 594
Db   1864 TACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 1923

Qy    595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
Db   1924 GATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGT 1983
```

```
Qy    615 SerVal------------------LysSerIleLys---SerAspPheIleAspLysAla 627
Db   1984 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAG 2043

Qy    628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
Db   2044 CTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATG 2103

Qy    648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
Db   2104 AAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAAT 2163

Qy    668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
Db   2164 GGAGGCCTGCAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGA 2223

Qy    688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db   2224 ATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCC 2283

Qy    707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db   2284 CTTGGAATTATTTCTTTTATTACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGA 2343

Qy    727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db   2344 GAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTG 2403

Qy    747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
Db   2404 AGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGAT 2463

Qy    767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db   2464 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCA 2523

Qy    787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
Db   2524 AATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTC 2583

Qy    807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db   2584 TTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGA 2643

Qy    827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db   2644 CAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 2703

Qy    847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db   2704 GAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCT 2763

Qy    867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db   2764 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 2823

Qy    887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db   2824 TTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCC 2883

Qy    907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db   2884 TACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTT 2943

Qy    927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db   2944 GTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCC 3003

Qy    947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db   3004 AAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTAC 3063

Qy    967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db   3064 AGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTA 3123

Qy    987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db   3124 TTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAG 3183

Qy   1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
Db   3184 AAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAG 3243

Qy   1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
Db   3244 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATA 3303

Qy   1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
Db   3304 AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATC 3363

Qy   1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db   3364 CTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCA 3423

Qy   1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db   3424 CAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTG 3483

Qy   1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
Db   3484 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAA 3543

Qy   1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
Db   3544 CAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAA 3603

Qy   1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
Db   3604 GCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCC 3663

Qy   1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db   3664 AGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGAC 3723

Qy   1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
Db   3724 TTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTG 3783

Qy   1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db   3784 GCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3825

RESULT 13
AAH57442
ID   AAH57442 standard; cDNA; 4349 BP.
XX
AC   AAH57442;
XX
DT   10-SEP-2001  (first entry)
XX
DE   Human intestine cell specific cDNA sequence SEQ ID NO:282.
XX
KW   Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
KW   lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
KW   metabolic disease; developmental disease; cytostatic; immunomodulatory;
KW   neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
XX
OS   Homo sapiens.
XX
PN   WO200132927-A2.
XX
PD   10-MAY-2001.
XX
PF   02-NOV-2000; 2000WO-US30396.
```

```
XX
PR   04-NOV-1999;   99US-0163508.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Sornasse T,  Seilhamer JJ,  Watson GA;
XX
DR   WPI; 2001-291057/30.
XX
PT   New cell and tissue specific polynucleotides useful for diagnosis,
PT   prognosis or monitoring of treatments for disorders where the gene is
PT   associated with a cancer, immunopathology or neuropathology -
XX
PS   Claim 1; Page 207-208; 327pp; English.
XX
CC   AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
CC   sequences (I). (I) can have cytostatic, immunomodulatory and
CC   neuroprotective activities, and can be used in gene therapy. (I) and
CC   proteins (II) encoded by then are used in high throughput screening
CC   assays to select DNA molecules, RNA molecules, peptide nucleic acids,
CC   mimetics, peptides, proteins, agonists, antagonists, antibodies or
CC   their fragments, immunoglobulins, inhibitors, drug compounds and
CC   pharmaceutical agents. Expression of (I) in a sample indicates the
CC   differentiation of embryonic stem cells into a tissue selected from
CC   brain, heart, kidney, liver, lung, skeletal muscle or pancreatic
CC   tissues. (I) and (II) are used to produce an expression profile that
CC   defines a metabolic or developmental process, treatment, condition,
CC   disease or disorder. The gene profile can be used for diagnosis,
CC   prognosis or monitoring of treatments and for investigating a
CC   predisposition to a disorder where the gene is associated with a
CC   cancer, immunopathology or neuropathology.
XX
SQ   Sequence 4349 BP; 1304 A; 816 C; 1048 G; 1181 T; 0 other;

Alignment Scores:
Pred. No.:              0          Length:         4349
Score:                  3521.50    Matches:        681
Percent Similarity:     73.92%     Conservative:   246
Best Local Similarity:  54.31%     Mismatches:     264
Query Match:            57.18%     Indels:         63
DB:                     22         Gaps:           10

US-09-873-409-5 (1-1222) x AAH57442 (1-4349)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
             |||:::::|||  ||||||::::::::::::||||||  |||||||||||   |||
Db     276 ATGGTGGTGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 335

Qy      21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
             |||  ||||||||:::|||                          :::|||:::|||
Db     336 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 395

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                          ::::::      |||      ||||||
Db     396 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 455

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                   |||      ||||||  ||||||||||        |||:::        |||
Db     456 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 515

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
             ||||||:::|||  |||  :::  |||  ||||||      :::||||||||||||
Db     516 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 575

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
             ||||||||:::::::::  |||:::||||||||||||||||      |||:::||||||
Db     576 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 635

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
             ||||||||||||:::|||   |||:::   ||||||::::::  |||||||||||||||
Db     636 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 695

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
             :::   ||||||:::|||:::||||||   |||      ||||||     :::
Db     696 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 755

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
             |||||||||||||||||||||||   |||:::   ||||||::::::    |||||||||
Db     756 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 815

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
                :::::::::::   |||   |||   |||||||||   ||||||:::||||||
Db     816 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 875

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
             |||||||||||||||||||||::::::||||||||||||||||||         |||
Db     876 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 935

Qy     201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
             :::|||||||||:::
Db     936 AAGAAAGAACTTGAA--------------------------------------------- 950

Qy     221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                                |||||   :::||||||::::::
Db     951 ------------------------------------AGGTACAACAAAAATTTAGAAGAA 974

Qy     241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
             ||||||   |||||||||:::   |||   :::   :::||||:::||||||     :::
Db     975 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 1034

Qy     261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
                :::      :::|||   ||||||||||||||||||||:::|||:::|||:::
Db    1035 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 1094

Qy     281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
             ||||||      |||:::||||||  ||||||   |||||||||||||||:::
Db    1095 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1148

Qy     301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
             ::::::   :::|||   |||   |||      |||   ||||||   ||||||||||||
Db    1149 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1208

Qy     321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
             |||:::   ||||||:::::::|||||   |||||||||||||:::::::|||   :::
Db    1209 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1268

Qy     341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
             |||:::||||||::::::|||:::|||     :::||||||:::||||||   |||:::|||
Db    1269 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1328

Qy     361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
             |||||||||      :::|||||||||||||||||||||||||:::::::::||||||:::
Db    1329 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1388

Qy     381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
             ||||||||||||||||||   :::|||   ||||||||||||   |||||||||:::|||
Db    1389 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1448

Qy     401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
             |||||||||||||||      :::|||   :::   ||||||     ||||||||:::
Db    1449 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1508

Qy     421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
             |||||||||   |||:::   ||||||||||||||||||||||||||||||||||||
Db    1509 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1568

Qy     441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
             |||||||||:::   ||||||:::||||||||:::::::||||||   :::|||:::|||
Db    1569 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1628

Qy     461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
             :::|||   :::||||||||||||||||||||||||:::   |||:::||||||:::
Db    1629 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1688

Qy     481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
             |||||||||||||||:::||||||||:::|||||||||||||||||||||||||||
```

```
Db   1689 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 1748

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   1749 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 1808

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db   1809 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACC 1868

Qy    541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db   1869 ACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC 1928

Qy    561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db   1929 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATT 1988

Qy    581 TyrTyrSerLeuValMetSerGln------------------AspIleLysLysAlaAsp 594
Db   1989 TACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 2048

Qy    595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
Db   2049 GATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGT 2108

Qy    615 SerVal------------------LysSerIleLys---SerAspPheIleAspLysAla 627
Db   2109 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAG 2168

Qy    628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
Db   2169 CTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATG 2228

Qy    648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
Db   2229 AAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAAT 2288

Qy    668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
Db   2289 GGAGGCCTGCAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGA 2348

Qy    688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db   2349 ATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCC 2408

Qy    707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db   2409 CTTGGAATTATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGA 2468

Qy    727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db   2469 GAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTG 2528

Qy    747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
Db   2529 AGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGAT 2588

Qy    767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db   2589 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCA 2648

Qy    787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
Db   2649 AATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTC 2708

Qy    807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db   2709 TTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGA 2768

Qy    827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db   2769 CAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 2828

Qy    847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db   2829 GAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCT 2888

Qy    867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db   2889 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 2948

Qy    887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db   2949 TTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCC 3008

Qy    907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db   3009 TACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTT 3068

Qy    927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db   3069 GTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCC 3128

Qy    947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db   3129 AAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTAC 3188

Qy    967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db   3189 AGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTA 3248

Qy    987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db   3249 TTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAG 3308

Qy   1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
Db   3309 AAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAG 3368

Qy   1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
Db   3369 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATA 3428

Qy   1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
Db   3429 AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATC 3488

Qy   1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db   3489 CTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCA 3548

Qy   1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db   3549 CAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTG 3608

Qy   1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
Db   3609 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAA 3668

Qy   1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
Db   3669 CAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAA 3728

Qy   1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
Db   3729 GCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCC 3788

Qy   1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db   3789 AGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGAC 3848

Qy   1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
Db   3849 TTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTG 3908
```

```
Qy   1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
               :::   |||||||||   :::|||:::   |||:::
Db   3909 GCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3950

RESULT 14
AAZ94738
ID   AAZ94738 standard; cDNA; 4646 BP.
XX
AC   AAZ94738;
XX
DT   01-AUG-2000  (first entry)
XX
DE   Human ATP binding cassette ABCB1 (MDR1) cDNA.
XX
KW   ABCB1; ATP binding cassette; human; cholesterol; lipid disorder;
KW   atherosclerosis; lipid disorder; dyslipidemia; psoriasis;
KW   lupus erythematosus; diagnosis; gene therapy; MDR1;
KW   multidrug resistance; chromosome 7q21; ss.
XX
OS   Homo sapiens.
XX
PN   WO200018912-A2.
XX
PD   06-APR-2000.
XX
PF   21-SEP-1999;   99WO-EP06991.
XX
PR   25-SEP-1998;   98US-0101706.
XX
PA   (FARB ) BAYER AG.
XX
PI   Schmitz G,  Klucken J;
XX
DR   WPI; 2000-293151/25.
XX
PT   Adenosine triphosphate binding proteins useful for identifying agents
PT   for treating atherosclerosis and other inflammatory disorders -
XX
PS   Claim 9; Page 110-112; 154pp; English.
XX
CC   The present sequence is that of human ATP binding cassette
CC   subfamily B protein ABCB1 cDNA.  The cDNA was identified using a
CC   differential display method in which monocytes from peripheral
CC   blood were subjected to macrophage differentiation and cholesterol
CC   loading with acetylated low density lipoproteins and subsequent
CC   deloading with high density lipoprotein (HDL3) to identify
CC   cholesterol sensitive genes.  The gene maps to chromosome 7q21
CC   and is also termed MDR1 (multidrug resistance).  The invention
CC   provides cholesterol-sensitive ABC genes (see AAZ94734-63).  These
CC   genes, and polypeptides encoded by them, can be used for diagnostic
CC   and therapeutic applications, and for biochemical or cell-based
CC   assays to screen for pharmacologically active modulator compounds
CC   useful for the treatment of lipid disorders, atherosclerosis or
CC   other inflammatory diseases such as psoriasis and lupus
CC   erythematosus.
XX
SQ   Sequence 4646 BP; 1371 A; 892 C; 1129 G; 1254 T; 0 other;

Alignment Scores:
Pred. No.:              0            Length:        4646
Score:                  3521.50      Matches:       681
Percent Similarity:     73.92%       Conservative:  246
Best Local Similarity:  54.31%       Mismatches:    264
Query Match:            57.18%       Indels:        63
DB:                     21           Gaps:          10

US-09-873-409-5 (1-1222) x AAZ94738 (1-4646)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          |||:::::|||   |||||:::::::::::||||||   |||||||||||    |||
Db    575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy     21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
          |||   |||||||||:::|||                          :::|||:::|||
Db    635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe----------------- 43
                :::::          |||      ||||||
Db    695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                |||        ||||||  |||||||||         |||:::       |||
Db    755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
          ||||||:::|||   |||   :::   |||   ||||||      :::||||||||||||
Db    815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
          |||||||||:::::::::   |||:::||||||||||||||||       |||:::||||||
Db    875 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
          ||||||||||||:::|||   |||:::   ||||||::::::   |||||||||||||||
Db    935 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 994

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
            :::   ||||||:::|||:::||||||       |||      ||||||      :::
Db    995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
          |||||||||||||||||||||   |||:::   ||||||::::::      ||||||||||
Db   1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
                ::::::::::::   |||   |||   ||||||||||   ||||||:::||||||
Db   1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
          |||||||||||||||||||||||||:::::::||||||||||||||||||      |||
Db   1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
          :::|||||||||:::
Db   1235 AAGAAAGAACTTGAA--------------------------------------------- 1249

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                             ||||||   :::||||||::::::
Db   1250 -----------------------------------AGGTACAACAAAAATTTAGAAGAA 1273

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
          ||||||      |||||||||:::   |||   :::   :::|||:::||||||   :::
Db   1274 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 1333

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
                :::      :::|||   ||||||||||||||||||||:::|||:::|||:::
Db   1334 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 1393

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
          ||||||      |||:::||||||   ||||||   ||||||||||||||||:::
Db   1394 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1447

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
          ::::::   :::|||   |||   |||      |||   ||||||   ||||||||||||
Db   1448 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1507

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
          |||:::   ||||||::::::||||||   ||||||||||||||||:::::::|||   :::
Db   1508 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1567

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
          |||:::||||||::::::|||:::|||   :::||||||:::||||||   |||:::|||
Db   1568 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1627

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
          |||||||||   :::|||||||||||||||||||||||||:::::::::||||||:::
```

```
Db   1628 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1687

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   1688 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1747

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   1748 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1807

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   1808 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1867

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   1868 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1927

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   1928 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1987

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   1988 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 2047

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   2048 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 2107

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db   2108 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACC 2167

Qy    541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db   2168 ACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC 2227

Qy    561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db   2228 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATT 2287

Qy    581 TyrTyrSerLeuValMetSerGln------------------AspIleLysLysAlaAsp 594
Db   2288 TACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 2347

Qy    595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
Db   2348 GATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGT 2407

Qy    615 SerVal------------------LysSerIleLys---SerAspPheIleAspLysAla 627
Db   2408 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAG 2467

Qy    628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
Db   2468 CTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATG 2527

Qy    648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
Db   2528 AAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAAT 2587

Qy    668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
Db   2588 GGAGGCCTGCAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGA 2647

Qy    688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db   2648 ATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCC 2707

Qy    707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db   2708 CTTGGAATTATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGA 2767

Qy    727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db   2768 GAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTG 2827

Qy    747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
Db   2828 AGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGAT 2887

Qy    767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db   2888 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCA 2947

Qy    787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
Db   2948 AATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTC 3007

Qy    807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db   3008 TTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGA 3067

Qy    827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db   3068 CAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 3127

Qy    847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db   3128 GAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCT 3187

Qy    867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db   3188 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 3247

Qy    887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db   3248 TTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCC 3307

Qy    907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db   3308 TACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTT 3367

Qy    927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db   3368 GTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCC 3427

Qy    947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db   3428 AAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTAC 3487

Qy    967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db   3488 AGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTA 3547

Qy    987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db   3548 TTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAG 3607

Qy   1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
Db   3608 AAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAG 3667

Qy   1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
Db   3668 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATA 3727

Qy   1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
Db   3728 AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATC 3787

Qy   1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db   3788 CTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCA 3847
```

```
Qy     1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
                 :::|||||       ||||||      |||||||||||:::||||||||    |||
Db     3848 CAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTG 3907

Qy     1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
            |||   ||||||:::|||:::||||||     ||||||   |||||||||||||||||||
Db     3908 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAA 3967

Qy     1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
            ||||||:::||||||||||||||||||:::::::::|||   ||||||||||||||||||
Db     3968 CAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAA 4027

Qy     1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
            ||||||||||||||||||   :::||||||||||||||||||   |||||||||||||||
Db     4028 GCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCC 4087

Qy     1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
            |||   ||||||||||||:::|||:::   ||||||||||||   |||||||||||||||
Db     4088 AGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGAC 4147

Qy     1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
            ||||||||||||   ||||||:::::||||||   ||||||||||||:::||||||
Db     4148 TTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTG 4207

Qy     1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
                  :::   |||||||||   :::|||:::   |||:::
Db     4208 GCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 4249

RESULT 15
AAD38994
ID   AAD38994 standard; DNA; 4646 BP.
XX
AC   AAD38994;
XX
DT   23-SEP-2002  (first entry)
XX
DE   Human mdr1 gene.
XX
KW   Human; haematologic malignancy; multidrug resistance; MDR; SUMO-1;
KW   hypoxia inducible factor-1; small ubiquitin-like-modifier; HIF-1;
KW   lymphoid disorder; chronic lymphoproliferative disorder; lymphoma;
KW   myeloid disorder; lymphocytic leukaemia; thrombocythaemia; myeloma;
KW   angiogenic myeloid metaplasia; myeloid leukaemia; gene therapy;
KW   polycythaemia vera; hypoxia responsive element; HRE; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             425..4267
FT                   /*tag=  a
FT                   /product= "Human MDR protein"
XX
PN   WO200234291-A2.
XX
PD   02-MAY-2002.
XX
PF   25-OCT-2001; 2001WO-US49856.
XX
PR   26-OCT-2000; 2000US-243542P.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI   Colgan SP;
XX
DR   WPI; 2002-471427/50.
DR   P-PSDB; AAE24211.
XX
PT   Treating a subject (at risk of) having a hematologic malignancy or
PT   multidrug resistance, e.g. lymphoma or myeloma, by administering
PT   hypoxia inducible factor 1 binding molecules or small
PT   ubiquitin-like-modifier-1 binding molecules  -
XX
PS   Disclosure; Page 53-55; 92pp; English.
XX
CC   The invention relates to a method of treating a subject having or at
CC   risk of developing a haematologic malignancy or multidrug resistance
CC   (MDR). The method involves administering hypoxia inducible factor-1
CC   (HIF-1) binding molecules or small ubiquitin-like-modifier (SUMO)-1
CC   binding molecules or HIF-1-SUMO-1 complex modulators. mdr1-hypoxia
CC   responsive element (HRE) binding molecules or antisense nucleic
CC   acid molecules and SUMO-1 binding molecules or antisense molecules
CC   are useful for treating a subject having or at risk of developing
CC   haematologic malignancy or MDR (e.g. a lymphoid disorder or a myeloid
CC   disorder). The lymphoid disorders include lymphocytic leukaemia or
CC   chronic lymphoproliferative disorders e.g. lymphoma, myeloma or chronic
CC   lymphoid leukaemia. The myeloid disorders include chronic or acute
CC   myeloid leukaemia, e.g. angiogenic myeloid metaplasia, essential
CC   thrombocythaemia or polycythaemia vera. The invention is used in gene
CC   therapy. The present sequence is human mdr1 gene.
XX
SQ   Sequence 4646 BP; 1371 A; 892 C; 1129 G; 1254 T; 0 other;

Alignment Scores:
Pred. No.:               0              Length:         4646
Score:                   3521.50        Matches:        681
Percent Similarity:      73.92%         Conservative:   246
Best Local Similarity:   54.31%         Mismatches:     264
Query Match:             57.18%         Indels:         63
DB:                      24             Gaps:           10

US-09-873-409-5 (1-1222) x AAD38994 (1-4646)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            |||::::::|||   ||||||::::::::::::||||||   ||||||||||||   |||
Db      575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy       21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
            |||   |||||||||:::|||                           :::|||:::|||
Db      635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy       32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                  ::::::      |||         ||||||
Db      695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy       44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                  |||         ||||||   |||||||||         |||:::      |||
Db      755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy       62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
            ||||||:::|||   |||   :::   |||         :::||||||||||||
Db      815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy       82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
            |||||||||:::::::::   |||:::||||||||||||||||||   |||:::||||||
Db      875 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy      102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
            ||||||||||||:::|||   |||:::   |||||||||::::::   |||||||||||||
Db      935 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 994

Qy      121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
               :::   ||||||:::|||:::||||||      |||      ||||||      :::
Db      995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy      141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
            ||||||||||||||||||||||   |||:::   ||||||::::::      |||||||||
Db     1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy      161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
               ::::::::::::   |||   |||   |||||||||   ||||||:::||||||
Db     1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy      181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
            |||||||||||||||||||||||::::::||||||||||||||||||||      |||
Db     1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234
```

```
Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db   1235 AAGAAAGAACTTGAA--------------------------------------------- 1249

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db   1250 -----------------------------------AGGTACAACAAAAATTTAGAAGAA 1273

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db   1274 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 1333

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db   1334 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 1393

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db   1394 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1447

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db   1448 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1507

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   1508 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1567

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   1568 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1627

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   1628 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1687

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   1688 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1747

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   1748 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1807

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   1808 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1867

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   1868 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1927

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   1928 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1987

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   1988 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 2047

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   2048 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 2107

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db   2108 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACC 2167

Qy    541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db   2168 ACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC 2227

Qy    561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db   2228 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATT 2287

Qy    581 TyrTyrSerLeuValMetSerGln------------------AspIleLysLysAlaAsp 594
Db   2288 TACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 2347

Qy    595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
Db   2348 GATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGT 2407

Qy    615 SerVal------------------LysSerIleLys---SerAspPheIleAspLysAla 627
Db   2408 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAG 2467

Qy    628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
Db   2468 CTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATG 2527

Qy    648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
Db   2528 AAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAAT 2587

Qy    668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
Db   2588 GGAGGCCTGCAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGA 2647

Qy    688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db   2648 ATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCC 2707

Qy    707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db   2708 CTTGGAATTATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGA 2767

Qy    727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db   2768 GAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTG 2827

Qy    747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
Db   2828 AGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGAT 2887

Qy    767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db   2888 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCA 2947

Qy    787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
Db   2948 AATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTC 3007

Qy    807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db   3008 TTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGA 3067

Qy    827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db   3068 CAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 3127

Qy    847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db   3128 GAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCT 3187

Qy    867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db   3188 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 3247

Qy    887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db   3248 TTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCC 3307

Qy    907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
```

```
Db   3308 TACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTT 3367

Qy    927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
             :::|||||||||||:::|||:::                 ||||||:::|||:::||||||
Db   3368 GTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCC 3427

Qy    947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
          |||      |||||||||:::         ::::::||||||   |||    ||||||||
Db   3428 AAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTAC 3487

Qy    967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
          |||   ||||||         |||:::|||   ||||||||||:::   |||   ||||||
Db   3488 AGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTA 3547

Qy    987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
          |||   ||||||   |||||||||:::      :::|||:::|||||||||||    ::::::
Db   3548 TTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAG 3607

Qy   1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
          :::|||:::|||:::|||   |||||||||||||||||||||||||||||||   ||||||
Db   3608 AAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAG 3667

Qy   1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
          ||||||:::|||   |||||||||:::   |||:::||||||   ||||||    :::
Db   3668 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATA 3727

Qy   1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
          |||   ||||||||||||||||||:::    :::   ||||||   |||||||||:::
Db   3728 AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATC 3787

Qy   1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
          ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||
Db   3788 CTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCA 3847

Qy   1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
             :::||||||      ||||||      ||||||||||||:::|||||||||    |||
Db   3848 CAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTG 3907

Qy   1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
          |||   ||||||:::|||:::||||||   ||||||   ||||||||||||||||||||
Db   3908 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAA 3967

Qy   1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
          ||||||:::|||||||||||||||||||:::::::::|||   |||||||||||||||||
Db   3968 CAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAA 4027

Qy   1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
          ||||||||||||||||||     :::||||||||||||||||   |||||||||||||||
Db   4028 GCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCC 4087

Qy   1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
          |||   ||||||||||||:::|||:::   ||||||||||||   ||||||||||||||||
Db   4088 AGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGAC 4147

Qy   1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
          ||||||||||||      ||||||::::::||||||   ||||||||||||:::||||||
Db   4148 TTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTG 4207

Qy   1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
                :::   |||||||||   :::|||:::   |||:::
Db   4208 GCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 4249
```

Search completed: March 30, 2003, 03:16:38
Job time : 1426.3 secs

GenCore version 5.1.4_p5_4578


```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 30, 2003, 01:24:25 ; Search time 16520.1 Seconds
                                           (without alignments)
                                           2152.747 Million cell updates/sec

Title:          US-09-873-409-5
Perfect score:  6159
Sequence:       1 MILGILASLVNGACLPLMPL..........QELLRNRDIYFKLVNAQSVQ 1222

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters:      4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115418_19228/app_query.fasta_1.7544
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_54887_@runat_27032003_115418_19228 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      GenEmbl:*
                1:  gb_ba:*
                2:  gb_htg:*
                3:  gb_in:*
                4:  gb_om:*
                5:  gb_ov:*
                6:  gb_pat:*
                7:  gb_ph:*
                8:  gb_pl:*
                9:  gb_pr:*
                10:  gb_ro:*
                11:  gb_sts:*
                12:  gb_sy:*
                13:  gb_un:*
                14:  gb_vi:*
                15:  em_ba:*
                16:  em_fun:*
                17:  em_hum:*
                18:  em_in:*
                19:  em_mu:*
                20:  em_om:*
                21:  em_or:*
                22:  em_ov:*
                23:  em_pat:*
                24:  em_ph:*
                25:  em_pl:*
                26:  em_ro:*
                27:  em_sts:*
                28:  em_un:*
                29:  em_vi:*
                30:  em_htg_hum:*
                31:  em_htg_inv:*
                32:  em_htg_other:*
                33:  em_htg_mus:*
                34:  em_htg_pln:*
                35:  em_htg_rod:*
                36:  em_htg_mam:*
                37:  em_htg_vrt:*
                38:  em_sy:*
                39:  em_htgo_hum:*
                40:  em_htgo_mus:*
                41:  em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6157	100.0	3702	6	AX339031	AX339031 Sequence
2	5988.5	97.2	3621	6	AX339032	AX339032 Sequence
3	5737	93.1	3699	6	AX478104	AX478104 Sequence
4	5328	86.5	3177	6	AX339030	AX339030 Sequence
5	4561.5	74.1	2856	6	AX339028	AX339028 Sequence
6	3550.5	57.6	4296	10	CRUPGPI	M60040 C.griseus P
7	3549.5	57.6	4304	10	CRUPGP1165	M59253 Chinese ham
8	3533.5	57.4	4279	6	AX105082	AX105082 Sequence
9	3533	57.4	4195	6	AX108656	AX108656 Sequence
10	3532	57.3	4927	10	AF257746	AF257746 Rattus no
11	3530.5	57.3	3860	6	AX322787	AX322787 Sequence
12	3530.5	57.3	4279	6	AX105078	AX105078 Sequence
13	3530.5	57.3	4279	6	AX105080	AX105080 Sequence
14	3530.5	57.3	4646	6	AX391099	AX391099 Sequence
15	3528.5	57.3	4045	12	AF269224	AF269224 Synthetic
16	3527.5	57.3	4378	6	E02326	E02326 Multidrug r
17	3526	57.2	3858	4	OAU78609	U78609 Ovis aries
18	3525.5	57.2	4279	6	AX105057	AX105057 Sequence
19	3525.5	57.2	4317	4	AF045016	AF045016 Canis fam
20	3525.5	57.2	4317	6	AX105059	AX105059 Sequence
21	3524.5	57.2	4186	6	AX108654	AX108654 Sequence
22	3521.5	57.2	3860	6	AX322789	AX322789 Sequence
23	3521.5	57.2	4646	6	AX336420	AX336420 Sequence
24	3521.5	57.2	4646	6	AX336708	AX336708 Sequence
25	3521.5	57.2	4646	6	I49610	I49610 Sequence 2
26	3521.5	57.2	4646	9	HUMMDR1	M14758 Homo sapien
27	3521.5	57.2	4669	6	I08557	I08557 Sequence 3
28	3521.5	57.2	6505	6	AR028671	AR028671 Sequence
29	3521.5	57.2	8630	6	AX012320	AX012320 Sequence
30	3521.5	57.2	8630	6	AX012321	AX012321 Sequence
31	3521.5	57.2	9318	6	AR028672	AR028672 Sequence
32	3520	57.2	4264	6	AR051647	AR051647 Sequence
33	3520	57.2	4264	6	AR051650	AR051650 Sequence
34	3518	57.1	4323	10	AF286167	AF286167 Rattus no
35	3514	57.1	3840	6	AX481416	AX481416 Sequence
36	3514	57.1	3988	6	AX024454	AX024454 Sequence
37	3514	57.1	4192	9	AF016535	AF016535 Homo sapi
38	3505.5	56.9	4280	10	CRUPGPII	M60041 C.griseus P
39	3503.5	56.9	4669	6	AR091275	AR091275 Sequence
40	3503.5	56.9	4669	6	AR203322	AR203322 Sequence
41	3503	56.9	4189	6	AX322791	AX322791 Sequence
42	3503	56.9	4298	10	MUSMDR	M14757 Mouse multi
43	3499.5	56.8	3934	4	CFA419568	AJ419568 Canis fam
44	3499	56.8	3888	10	AY082609	AY082609 Rattus no
45	3499	56.8	4788	6	AX322793	AX322793 Sequence

ALIGNMENTS

RESULT 1

```
AX339031
LOCUS       AX339031                3702 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 13 from Patent WO0194400.
ACCESSION   AX339031
VERSION     AX339031.1  GI:18129123
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 13 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..3702
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1098 a    708 c    856 g   1039 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:               0              Length:        3702
Score:                   6157.00        Matches:       1222
Percent Similarity:      100.00%        Conservative:  0
Best Local Similarity:   100.00%        Mismatches:    0
Query Match:             99.97%         Indels:        0
DB:                      6              Gaps:          0

US-09-873-409-5 (1-1222) x AX339031 (1-3702)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       34 ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 93

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153

Qy       41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      154 TCTTTCTTCAGGTTGACCCTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGT 213

Qy       61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      214 TACATACAGATTTCCTTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAA 273

Qy       81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      274 CAGTTTTTTCATTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333

Qy      101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      334 GAACTTAACACTCGCATGACAGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATT 393

Qy      121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAG 453

Qy      141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      454 GGCTGGAAACTCACCCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA 513

Qy      161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573

Qy      181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAG 633

Qy      201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      634 GAGAAAGAACTTCAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTTGGTTTTATTTT 693

Qy      221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      694 CCCCAGTGGCTACTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGAATCTCAAAGAT 753

Qy      241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      754 GCAAAGGATTTTGGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTAC 813

Qy      261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      814 TTCTTTATGAATGGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAAT 873

Qy      281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      874 GGAGAACCTGGATATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGC 933

Qy      301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      934 AGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCT 993

Qy      321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      994 GCCTTTCATATTTTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCT 1053

Qy      341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1054 GGATATAAACCTGAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTAT 1113

Qy      361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1114 CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAG 1173

Qy      381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1174 ACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG 1233

Qy      401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1234 AGGTTATATGATCCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1293

Qy      421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1294 AATGTGCGGCATTATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGG 1353

Qy      441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1354 ACCACCATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1413

Qy      461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1414 AGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAAT 1473

Qy      481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1474 ACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATT 1533

Qy      501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1534 GCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTG 1593

Qy      521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1594 GATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTGGAGAAGGCGAGCAAAGGTCGGACT 1653

Qy      541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1654 ACAATCGTGGTAGCACACCGACTTTCTACTATTCGAAGTGCAGATTTGATTGTGACCCTA 1713

Qy      561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
```

```
                ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1714 AAGGATGGAATGCTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTA 1773

Qy        581 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1774 TATTATTCACTTGTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATG 1833

Qy        601 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1834 ACATATTCTACTGAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAG 1893

Qy        621 SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGlu 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1894 TCAGACTTCATTGACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAA 1953

Qy        641 ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGly 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1954 GTCTCTCTATTAAAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGG 2013

Qy        661 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2014 ACATTGGCTTCTGTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAA 2073

Qy        681 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2074 ATTATAACCATGTTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTAT 2133

Qy        701 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2134 TCCATGATATTCGTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTA 2193

Qy        721 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2194 TTTTACGGCAGAGCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCC 2253

Qy        741 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2254 ATGTTATATCAGGATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACA 2313

Qy        761 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2314 ACAATATTAGCCATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTC 2373

Qy        781 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrp 800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2374 TTAACACAAAATGCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGG 2433

Qy        801 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2434 GAGATGACATTCCTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAA 2493

Qy        821 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2494 ACCGCAGCAATGACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAG 2553

Qy        841 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2554 ATAGCAACTGAAGCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCC 2613

Qy        861 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2614 TTCGAGCAAATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCA 2673

Qy        881 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2674 CAGATTATTGGAAGCTGTTATGCATTCAGCCATGCCTTTATATATTTTGCCTATGCAGCA 2733

Qy        901 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2734 GGGTTTCGATTTGGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTC 2793

Qy        921 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2794 ATAGTTTTTACTGCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCT 2853

Qy        941 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2854 CCTGAATATTCCAAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAA 2913

Qy        961 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2914 CCAAATATAGACAGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTA 2973

Qy        981 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2974 GAGTTTCGAGAAGTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGC 3033

Qy       1001 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3034 TTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGG 3093

Qy       1021 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3094 AAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTG 3153

Qy       1041 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3154 TTTGATGGTGTGGATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATC 3213

Qy       1061 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3214 GTTCCTCAAGAGCCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGAC 3273

Qy       1081 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3274 AACAGCCGTGTGGTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCAT 3333

Qy       1101 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 1120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3334 TCTTTTATTGAAGGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAG 3393

Qy       1121 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3394 CTTTCTGGCGGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAA 3453

Qy       1141 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3454 ATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAG 3513

Qy       1161 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3514 CATGCCCTTGATAAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCT 3573

Qy       1181 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3574 GCAATTCAGAACGCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGA 3633

Qy       1201 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3634 ACTCATCAAGAGCTCCTGAGAAATCGAGACATATATTTAAGTTAGTGAATGCACAGTCA 3693

Qy       1221 ValGln 1222
              ||||||
Db       3694 GTGCAG 3699
```

```
RESULT 2
AX339032
LOCUS       AX339032                3621 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 14 from Patent WO0194400.
ACCESSION   AX339032
VERSION     AX339032.1  GI:18129124
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
REFERENCE    1
  AUTHORS    Frank,M.H. and Sayegh,M.H.
  TITLE      A gene encoding a multidrug resistance human p-glycoprotein
             homologue on chromosome 7p15-21 and uses thereof
  JOURNAL    Patent: WO 0194400-A 14 13-DEC-2001;
             THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES              Location/Qualifiers
     source           1..3621
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
BASE COUNT     1081 a    696 c    842 g   1002 t
ORIGIN

Alignment Scores:
Pred. No.:               0            Length:         3621
Score:                   5988.50      Matches:        1195
Percent Similarity:      97.79%       Conservative:   0
Best Local Similarity:   97.79%       Mismatches:     0
Query Match:             97.23%       Indels:         27
DB:                      6            Gaps:           1

US-09-873-409-5 (1-1222) x AX339032 (1-3621)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      34 ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 93

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153

Qy      41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     154 TCTTTCTTCAGGTTGACCCTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGT 213

Qy      61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     214 TACATACAGATTTCCTTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAA 273

Qy      81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     274 CAGTTTTTTCATTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333

Qy     101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     334 GAACTTAACACTCGCATGACAGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATT 393

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAG 453

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     454 GGCTGGAAACTCACCCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA 513

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAG 633

Qy     201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
           |||||||||||||||
Db     634 GAGAAAGAACTTCAA--------------------------------------------- 648

Qy     221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                                    ||||||||||||||||||||||||
Db     649 ------------------------------------AGGTATACACAGAATCTCAAAGAT 672

Qy     241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     673 GCAAAGGATTTTGGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTAC 732

Qy     261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     733 TTCTTTATGAATGGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAAT 792

Qy     281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     793 GGAGAACCTGGATATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGC 852

Qy     301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     853 AGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCT 912

Qy     321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     913 GCCTTTCATATTTTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCT 972

Qy     341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     973 GGATATAAACCTGAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTAT 1032

Qy     361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1033 CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAG 1092

Qy     381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1093 ACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG 1152

Qy     401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1153 AGGTTATATGATCCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1212

Qy     421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1213 AATGTGCGGCATTATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGG 1272

Qy     441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1273 ACCACCATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1332

Qy     461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1333 AGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAAT 1392

Qy     481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1393 ACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATT 1452

Qy     501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1453 GCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTG 1512

Qy     521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1513 GATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTGGAGAAGGCGAGCAAAGGTCGGACT 1572

Qy     541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1573 ACAATCGTGGTAGCACACCGACTTTCTACTATTCGAAGTGCAGATTTGATTGTGACCCTA 1632

Qy     561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1633 AAGGATGGAATGCTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTA 1692

Qy     581 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1693 TATTATTCACTTGTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATG 1752

Qy     601 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 620
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1753 ACATATTCTACTGAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAG 1812
```

```
Qy     621 SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGlu 640
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1813 TCAGACTTCATTGACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAA 1872

Qy     641 ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGly 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1873 GTCTCTCTATTAAAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGG 1932

Qy     661 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 680
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1933 ACATTGGCTTCTGTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAA 1992

Qy     681 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 700
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1993 ATTATAACCATGTTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTAT 2052

Qy     701 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2053 TCCATGATATTCGTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTA 2112

Qy     721 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 740
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2113 TTTTACGGCAGAGCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCC 2172

Qy     741 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 760
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2173 ATGTTATATCAGGATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACA 2232

Qy     761 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2233 ACAATATTAGCCATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTC 2292

Qy     781 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrp 800
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2293 TTAACACAAAATGCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGG 2352

Qy     801 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 820
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2353 GAGATGACATTCCTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAA 2412

Qy     821 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2413 ACCGCAGCAATGACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAG 2472

Qy     841 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 860
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2473 ATAGCAACTGAAGCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCC 2532

Qy     861 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 880
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2533 TTCGAGCAAATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCA 2592

Qy     881 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 900
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2593 CAGATTATTGGAAGCTGTTATGCATTCAGCCATGCCTTTATATATTTTGCCTATGCAGCA 2652

Qy     901 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 920
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2653 GGGTTTCGATTTGGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTC 2712

Qy     921 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 940
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2713 ATAGTTTTTACTGCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCT 2772

Qy     941 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 960
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2773 CCTGAATATTCCAAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAA 2832

Qy     961 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 980
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2833 CCAAATATAGACAGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTA 2892

Qy     981 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 1000
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2893 GAGTTTCGAGAAGTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGC 2952

Qy    1001 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 1020
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2953 TTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGG 3012

Qy    1021 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 1040
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3013 AAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTG 3072

Qy    1041 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 1060
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3073 TTTGATGGTGTGGATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATC 3132

Qy    1061 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 1080
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3133 GTTCCTCAAGAGCCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGAC 3192

Qy    1081 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 1100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3193 AACAGCCGTGTGGTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCAT 3252

Qy    1101 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 1120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3253 TCTTTTATTGAAGGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAG 3312

Qy    1121 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 1140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3313 CTTTCTGGCGGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAA 3372

Qy    1141 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 1160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3373 ATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAG 3432

Qy    1161 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 1180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3433 CATGCCCTTGATAAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCT 3492

Qy    1181 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 1200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3493 GCAATTCAGAACGCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGA 3552

Qy    1201 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3553 ACTCATCAAGAGCTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCA 3612

Qy    1221 ValGln 1222
           ||||||
Db    3613 GTGCAG 3618
```

```
RESULT 3
AX478104
LOCUS       AX478104                3699 bp    DNA     linear   PAT 12-AUG-2002
DEFINITION  Sequence 34 from Patent WO0240541.
ACCESSION   AX478104
VERSION     AX478104.1  GI:22217064
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Tang,Y.T., Yue,H., Nguyen,D.B., Hafalia,A.J., Elliott,V.S., Lu,Y.,
            Walia,N.K., Yao,M.G., Baughn,M.R., Gandhi,A.R., Ding,L.,
            Sanjanwala,M., Ramkumar,J., Arvizu,C., Gietzen,K.J., Lal,P.G.,
            Azimzai,Y., Khan,F.A., Thangavelu,K., Thornton,M., Lu,D.A.,
            Tribouley,C.M., Warren,B.A., Ison,C.H., Das,D., Raumann,B.E.,
            Policky,J.L. and Kearney,L.
  TITLE     Transporters and ion channels
  JOURNAL   Patent: WO 0240541-A 34 23-MAY-2002;
            Incyte Genomics, Inc. (US)
```

```
FEATURES             Location/Qualifiers
     source          1. .3699
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 7472030CB1"
BASE COUNT     1116 a    707 c    860 g   1016 t
ORIGIN

Alignment Scores:
Pred. No.:                0              Length:          3699
Score:                    5737.00        Matches:         1156
Percent Similarity:       93.78%         Conservative:    4
Best Local Similarity:    93.45%         Mismatches:      9
Query Match:              93.15%         Indels:          68
DB:                       6              Gaps:            4

US-09-873-409-5 (1-1222) x AX478104 (1-3699)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          |||||||||||||||   ||||||   |||||||||||||||||||||||||||||||||
Db    145 ATGATCCTGGGTATACTGACATCACTGTTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 204

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
             :::||||||||||||||||||||||||||||||||||||||||   |||||||||
Db    205 TGTATAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCACACTAACACAACA 264

Qy     41 SerPhePheArg------------------------------------------LeuThr 46
          ::::::                                                |||||
Db    265 AATTATCAGAACTGTACTCAGTCTCAAGAGAAGCTGAATGAAGATATGACTCTGTTGACC 324

Qy     47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    325 CTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGTTACATACAGATTTCCTTG 384

Qy     67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    385 TGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAACAGTTTTTCATTCAGTT 444

Qy     87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    445 TTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGTGAACTTAACACTCGCATG 504

Qy    107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
          |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    505 ACAGATGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATTGCTCTGTTGTTTCAA 564

Qy    126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    565 AACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAGGGCTGGAAACTCACC 624

Qy    146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    625 CTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCAGCATGTTCTAGGATG 684

Qy    166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    685 GTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCAGAA 744

Qy    186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    745 GAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAGGAGAAAGAACTTCAA 804

Qy    206 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeu 225

Db    804 ------------------------------------------------------------ 804

Qy    226 SerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 245
                                 ||||||||||||||||||||||||||||||||||||||||
Db    805 --------------------AGGTATACACAGAATCTCAAAGATGCAAAGGATTTTGGC 843

Qy    246 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 265
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    844 ATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAATGGA 903

Qy    266 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 285
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    904 ACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGATAT 963

Qy    286 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    964 ACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGTAGTTATTGCATTGGA 1023

Qy    306 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 325
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1024 GCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTC 1083

Qy    326 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 345
          |||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db   1084 CAGGTTATTGATAAGAAACCCAGTATAGGTAACTTTTCCACAGCTGGATATAAACCTGAA 1143

Qy    346 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 365
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1144 TCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGGCCATCT 1203

Qy    366 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 385
          |||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db   1204 ATCAAGATTCTGAAAGGTCTGAATCTCGGAATTAAGTCTGGAGAGACAGTCGCCTTGGTC 1263

Qy    386 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 405
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1264 GGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCG 1323

Qy    406 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 425
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1324 GATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTAT 1383

Qy    426 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 445
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1384 CGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAAC 1443

Qy    446 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 465
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1444 AATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAA 1503

Qy    466 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 485
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1504 GCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAA 1563

Qy    486 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 505
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1564 AAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTT 1623

Qy    506 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 525
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1624 CGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAG 1683

Qy    526 SerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 545
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1684 TCAGCTGTTCAAGCTGCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTAGCA 1743

Qy    546 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 565
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1744 CACCGACTTTCTACTATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATGCTG 1803

Qy    566 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 585
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1804 GCGGAGAAAGGAGCGCATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTTGTG 1863

Qy    586 MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 605
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1864 ATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAA 1923

Qy    606 ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIleAsp 625
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1924 AGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGAC 1983
```

```
Qy    626 LysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLys 645
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1984 AAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTAAAA 2043

Qy    646 IleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerVal 665
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2044 ATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGACATTGGCTTCTGTT 2103

Qy    666 LeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPhe 685
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2104 CTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATGTTT 2163

Qy    686 GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheVal 705
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2164 GGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTCGTC 2223

Qy    706 IleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAla 725
          ||||||||||||||||||||||||||||||||||||
Db   2224 ATTTGGGTGTTATTTGCTTTGTCAGTTATTTCATG------------------------ 2259

Qy    726 GlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAsp 745
                                                                ||||||
Db   2260 ------------------------------------------------------CAGGAT 2265

Qy    746 IleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIle 765
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2266 ATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCCATA 2325

Qy    766 AspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAla 785
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2326 GATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAATGCA 2385

Qy    786 ThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeu 805
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2386 ACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTCCTG 2445

Qy    806 IleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThr 825
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2446 ATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATGACT 2505

Qy    826 GlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAla 845
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2506 GGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCT 2565

Qy    846 LeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyr 865
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2566 TTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATGTAT 2625

Qy    866 GluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySer 885
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2626 GAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGC 2685

Qy    886 CysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGly 905
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2686 TGTTATGCATTCAGCCATGCCTTTATATATTTTGCCTATGCGGCAGGGTTTCGATTTGGA 2745

Qy    906 AlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAla 925
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2746 GCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACTGCA 2805

Qy    926 IleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLys 945
          |||||||||||||||||||||||||||:::||||||||||||||||||||||||||||||
Db   2806 ATTGCATATGGAGCTATGGCCATCGGAGAAACGCTCGTTTTGGCTCCTGAATATTCCAAA 2865

Qy    946 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSer 965
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2866 GCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGACAGC 2925

Qy    966 ArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 985
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2926 CGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAAGTC 2985

Qy    986 SerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIle 1005
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2986 TCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGTATT 3045

Qy   1006 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 1025
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3046 GAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCTGTT 3105

Qy   1026 GlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAsp 1045
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3106 CAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTGGAT 3165

Qy   1046 AlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 1065
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3166 GCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAGCCT 3225

Qy   1066 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValVal 1085
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3226 GTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTGGTG 3285

Qy   1086 ProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGly 1105
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3286 CCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAAGGT 3345

Qy   1106 LeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGln 1125
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3346 CTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGCCAG 3405

Qy   1126 LysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAsp 1145
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3406 AAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTGGAT 3465

Qy   1146 GluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLys 1165
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3466 GAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGATAAA 3525

Qy   1166 AlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla 1185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3526 GCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCA 3585

Qy   1186 AspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1205
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3586 GATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAGCTC 3645

Qy   1206 LeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1222
          |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3646 CTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3696

RESULT 4
AX339030
LOCUS       AX339030                3177 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 12 from Patent WO0194400.
ACCESSION   AX339030
VERSION     AX339030.1  GI:18129122
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 12 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..3177
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      970 a    601 c    735 g    870 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:             0              Length:         3177
```

```
Score:                    5328.00     Matches:        1058
Percent Similarity:       100.00%     Conservative:   0
Best Local Similarity:    100.00%     Mismatches:     0
Query Match:              86.51%      Indels:         0
DB:                       6           Gaps:           0

US-09-873-409-5 (1-1222) x AX339030 (1-3177)

Qy    165 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 184
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCA 60

Qy    185 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 204
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAGGAGAAAGAACTT 120

Qy    205 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 224
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 CAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTTGGTTTTATTTCCCCAGTGGCTA 180

Qy    225 LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 244
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGAATCTCAAAGATGCAAAGGATTTT 240

Qy    245 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 264
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAAT 300

Qy    265 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 284
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGA 360

Qy    285 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 304
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATT 420

Qy    305 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 324
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATT 480

Qy    325 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 344
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCT 540

Qy    345 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 364
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCA 600

Qy    365 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 384
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTG 660

Qy    385 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 404
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 720

Qy    405 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 424
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCAT 780

Qy    425 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 444
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGT 840

Qy    445 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 464
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGG 900

Qy    465 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 484
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGG 960

Qy    485 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 504
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTA 1020

Qy    505 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 524
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGC 1080

Qy    525 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 544
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AAGTCAGCTGTTCAAGCTGCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTA 1140

Qy    545 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 564
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GCACACCGACTTTCTACTATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATG 1200

Qy    565 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 584
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTT 1260

Qy    585 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 604
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACT 1320

Qy    605 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 624
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATT 1380

Qy    625 AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 644
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTA 1440

Qy    645 LysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 664
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 AAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCT 1500

Qy    665 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 684
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATG 1560

Qy    685 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 704
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 TTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTC 1620

Qy    705 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 724
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGA 1680

Qy    725 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 744
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAG 1740

Qy    745 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 764
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCC 1800

Qy    765 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 784
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAAT 1860

Qy    785 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 804
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTC 1920

Qy    805 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 824
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATG 1980

Qy    825 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 844
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 ACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAA 2040

Qy    845 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 864
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   2041 GCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATG 2100

Qy    865 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 884
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGA 2160

Qy    885 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 904
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 AGCTGTTATGCATTCAGCCATGCCTTTATATATTTGCCTATGCAGCAGGGTTTCGATTT 2220

Qy    905 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 924
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 GGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACT 2280

Qy    925 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 944
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 GCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCC 2340

Qy    945 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 964
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 AAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGAC 2400

Qy    965 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 984
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 AGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAA 2460

Qy    985 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 1004
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 GTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGT 2520

Qy   1005 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 1024
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 ATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCT 2580

Qy   1025 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 1044
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 GTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTG 2640

Qy   1045 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 1064
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 GATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAG 2700

Qy   1065 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 1084
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 CCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTG 2760

Qy   1085 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 1104
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 GTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAA 2820

Qy   1105 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 1124
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 GGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGC 2880

Qy   1125 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 1144
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 CAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTG 2940

Qy   1145 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1164
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 GATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGAT 3000

Qy   1165 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1184
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 AAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAC 3060

Qy   1185 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1204
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 GCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAG 3120

Qy   1205 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1222
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 CTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3174
```

```
RESULT 5
AX339028
LOCUS       AX339028                2856 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 10 from Patent WO0194400.
ACCESSION   AX339028
VERSION     AX339028.1  GI:18129120
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 10 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..2856
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      888 a    540 c    652 g    776 t
ORIGIN

Alignment Scores:
Pred. No.:              2.27e-296      Length:         2856
Score:                  4561.50        Matches:        912
Percent Similarity:     97.96%         Conservative:   0
Best Local Similarity:  97.96%         Mismatches:     1
Query Match:            74.06%         Indels:         18
DB:                     6              Gaps:           1

US-09-873-409-5 (1-1222) x AX339028 (1-2856)

Qy    293 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHis-PheGl 312
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    114 GTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATTGGAGCAGCAGTCCCTCATTATT-- 171

Qy    312 uThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPr 332
                                                                |||||||||||
Db    172 -------------------------------------------------GATAAGAAACC 182

Qy    332 oSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGl 352
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    183 CAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGA 242

Qy    352 uPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLe 372
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    243 ATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCT 302

Qy    372 uAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLy 392
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    303 GAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAA 362

Qy    392 sSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVa 412
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    363 GAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGT 422

Qy    412 lAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVa 432
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    423 GGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTGGT 482

Qy    432 lSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAs 452
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    483 TAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGAGA 542

Qy    452 pAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIl 472
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    543 TGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTTAT 602

Qy    472 eMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGl 492
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     603 CATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGG 662

Qy     492 yGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIl 512
Db     663 AGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGAT 722

Qy     512 eLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLe 532
Db     723 TTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACT 782

Qy     532 uGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleAr 552
Db     783 GGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTAGCACACCGACTTTCTACTATTCG 842

Qy     552 gSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAl 572
Db     843 AAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATGCTGGCGGAGAAAGGAGCACATGC 902

Qy     572 aGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLy 592
Db     903 TGAACTAATGGCAAAACGAGGTCTATATTATTCACTTGTGATGTCACAGGATATTAAAAA 962

Qy     592 sAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPr 612
Db     963 AGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCACTTCC 1022

Qy     612 oLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGl 632
Db    1023 TCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCCACCCA 1082

Qy     632 nSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPr 652
Db    1083 ATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAACAAGCC 1142

Qy     652 oGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPr 672
Db    1143 TGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCTGTTCTAAATGGAACTGTTCATCC 1202

Qy     672 oValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrTh 692
Db    1203 AGTATTTTCCATCATCTTTGCAAAAATTATAACCATGTTTGGAAATAATGATAAAACCAC 1262

Qy     692 rLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPh 712
Db    1263 ATTAAAGCATGATGCAGAAATTTATTCCATGATATTCGTCATTTTGGGTGTTATTTGCTT 1322

Qy     712 eValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetAr 732
Db    1323 TGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGAGCAGGGGAAATTTTAACGATGAG 1382

Qy     732 gLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLy 752
Db    1383 ATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAGGATATTGCCTGGTTTGATGAAAA 1442

Qy     752 sGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGl 772
Db    1443 GGAAAACAGCACAGGAGGCTTGACAACAATATTAGCCATAGATATAGCACAAATTCAAGG 1502

Qy     772 yAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVa 792
Db    1503 AGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAATGCAACTAACATGGGACTTTCAGT 1562

Qy     792 lIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVa 812
Db    1563 TATCATTTCCTTTATATATGGATGGGAGATGACATTCCTGATTCTGAGTATTGCTCCAGT 1622

Qy     812 lLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLy 832
Db    1623 ACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATGACTGGATTTGCCAACAAAGATAA 1682

Qy     832 sGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIl 852
Db    1683 GCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAGAATATACGTACTAT 1742

Qy     852 eValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGl 872
Db    1743 AGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATGTATGAAGAGATGCTTCAGACTCA 1802

Qy     872 nHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAl 892
Db    1803 ACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTATGCATTCAGCCATGC 1862

Qy     892 aPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGl 912
Db    1863 CTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTTGGAGCCTATTTAATTCAAGCTGG 1922

Qy     912 yArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAl 932
Db    1923 ACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACTGCAATTGCATATGGAGCTATGGC 1982

Qy     932 aIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHi 952
Db    1983 CATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCCAAAGCCAAATCGGGGGCTGCGCA 2042

Qy     952 sLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLy 972
Db    2043 TCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGACAGCCGCAGTCAAGAAGGGAAAAA 2102

Qy     972 sProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysAr 992
Db    2103 GCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTCTTCTATCCATGTCG 2162

Qy     992 gProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAl 1012
Db    2163 CCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGC 2222

Qy    1012 aPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTy 1032
Db    2223 ATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTA 2282

Qy    1032 rAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGl 1052
Db    2283 TGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTGGATGCAAAAGAATTGAATGTACA 2342

Qy    1052 nTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIl 1072
Db    2343 GTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAGCCTGTGCTCTTCAACTGCAGCAT 2402

Qy    1072 eAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGl 1092
Db    2403 TGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTGGTGCCATTAGATGAGATCAAAGA 2462

Qy    1092 uAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnTh 1112
Db    2463 AGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAAGGTCTCCCTGAGAAATACAACAC 2522

Qy    1112 rGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAl 1132
Db    2523 ACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCTATTGC 2582

Qy    1132 aArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAs 1152
Db    2583 AAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGA 2642

Qy    1152 pAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCy 1172
Db    2643 TAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGATAAAGCCAGGACGGGAAGGACATG 2702

Qy    1172 sLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHi 1192
Db    2703 CCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGGTTCTGCA 2762

Qy    1192 sAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTy 1212
Db    2763 CAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAGCTCCTGAGAAATCGAGACATATA 2822
```

```
Qy   1212 rPheLysLeuValAsnAlaGlnSerValGln 1222
          ||||||||||||||||||||||||||||||
Db   2823 TTTTAAGTTAGTGAATGCACAGTCAGTGCAG 2853

RESULT 6
CRUPGPI
LOCUS       CRUPGPI                 4296 bp    mRNA    linear   ROD 22-JUN-1995
DEFINITION  C.griseus P-glycoprotein (isoform class I) mRNA, complete cds.
ACCESSION   M60040
VERSION     M60040.1  GI:191164
KEYWORDS    multidrug resistance; p-glycoprotein; transmembrane protein.
SOURCE      C.griseus adult liver and ovary, cDNA to mRNA.
  ORGANISM  Cricetulus griseus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
            Cricetulus.
REFERENCE   1  (bases 1 to 4296)
  AUTHORS   Endicott,J.A., Sarangi,F. and Ling,V.
  TITLE     Complete cDNA sequences encoding the Chinese hamster P-glycoprotein
            gene family
  JOURNAL   DNA Seq. 2 (2), 89-101 (1991)
  MEDLINE   92135896
   PUBMED   1685679
FEATURES             Location/Qualifiers
     source          1..4296
                     /organism="Cricetulus griseus"
                     /db_xref="taxon:10029"
                     /tissue_type="liver and ovary"
                     /dev_stage="adult"
     CDS             103..3933
                     /note="putative"
                     /codon_start=1
                     /product="p-glycoprotein isoform I"
                     /protein_id="AAA68883.1"
                     /db_xref="GI:191165"
                     /translation="MEFEEDFSGRKDKNFLKMGRKSKKEKKEKKPVVSVFTMFRYAGW
                     LDRLYMLVGTLAAIIHGVALPLMMLVFGDMTDSFASVGNIPTNATNNATQVNASDIFG
                     KLEEEMTTYAYYYTGIGAGVLIVAYIQVSFWCLAAGRQIHKIRQKFFHAIMNQEIGWF
                     DVHDVGELNTRLTDDVSKINEGIGDKIGMFFQAMATFFGGFIIGFTRGWKLTLVILAI
                     SPVLGLSAGIWAKILSSFTDKELQAYAKAGAVAEEVLAAIRTVIAFGGQKKELERYNN
                     NLEEAKRLGIKKAITANISMGAAFLLIYASYALAFWYGTSLVISKEYSIGQVLTVFFA
                     VLIGAFSIGQASPNIEAFANARGAAYEIFNIIDNKPSIDSFSKNGYKPDNIKGNLEFK
                     NIHFSYPSRKDVQILKGLNLKVQSGQTVALVGNSGCGKSTTVQLLQRLYDPTEGVVSI
                     DGQDIRTINVRYLREIIGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKEANAYD
                     FIMKLPHKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTESEAVV
                     QAALDKAREGRTTIVIAHRLSTVRNADIIAGFDGGVIVEQGNHEELMREKGIYFKLVM
                     TQTAGNEIELGNEVGESKNEIDNLDMSSKDSASSLIRRRSTRRSIRGPHDQDRKLSTK
                     EALDEDVPPISFWRILKLNSSEWPYFVVGIFCAIVNGALQPAFSIIFSKVVGVFTRNT
                     DDETKRHDSNLFSLLFLILGVISFITFFLQGFTFGKAGEILTKRLRYMVFKSMLRQDV
                     SWFDNPKNTTGALTTRLANDAGQVKGATGARLAVITQNIANLGTGIIISLIYGWQLTL
                     LLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGSGKIATEAIENFRTVVSLTREQKFE
                     NMYAQSLQIPYRNALKKAHVFGITFSFTQAMMYFSYAACFRFGAYLVARELMTFENVL
                     LVFSAIVFGAMAVGQVSSFAPDYAKAKVSASHIIMIIEKVPSIDSYSTGGLKPNTLEG
                     NVKFNEVVFNYPTRPDIPVLQGLNLEVKKGQTLALVGSSGCGKSTVVQLLERFYDPMA
                     GTVFLDGKEVNQLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSQDEIERAA
                     KEANIHQFIESLPDKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEATSALD
                     TESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVIQNGKVKEHGTHQQLLAQKG
                     IYFSMVSVQAGAKR"
BASE COUNT     1271 a    892 c   1029 g   1104 t
ORIGIN

Alignment Scores:
Pred. No.:              2.05e-228      Length:         4296
Score:                  3550.50        Matches:        679
Percent Similarity:     73.75%         Conservative:   248
Best Local Similarity:  54.02%         Mismatches:     259
Query Match:            57.65%         Indels:         71
DB:                     10             Gaps:           9

US-09-873-409-5 (1-1222) x CRUPGPI (1-4296)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          |||::::::|||   ||||||::::::::::::|||         ||||||||||||   |||
Db    250 ATGCTGGTGGGGACTCTCGCAGCTATTATCCATGGAGTTGCACTCCCGCTTATGATGCTG 309

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsn------ 38
          |||     |||:::|||:::|||:::         |||          :::   ||||||
Db    310 GTGTTTGGAGACATGACAGACAGCTTTGCAAGTGTAGGAAACATCCCAACAAATGCTACT 369

Qy     38 ------------------------------------------------------------ 38

Db    370 AATAATGCCACTCAAGTCAATGCTTCAGACATCTTTGGCAAACTGGAGGAAGAAATGACC 429

Qy     39 ThrTyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIle 58
          ||||||::::::                    ||||||   |||||||||         ||||||
Db    430 ACGTACGCCTAT---------------TACTACACAGGGATTGGTGCTGGTGTGCTTATA 474

Qy     59 PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIle 78
                |||||||||:::|||     |||    :::   |||   ||||||        :::|||
Db    475 GTTGCCTACATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATT 534

Qy     79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp 98
          |||::::::|||||||||:::::::::      |||:::||||||||||||||||||      |||
Db    535 AGACAGAAGTTTTTCATGCTATAATGAATCAGGAGATAGGCTGGTTTGACGTGCACGAC 594

Qy     99 IleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly 117
          :::||||||||||||||||||:::|||      |||:::   ||||||::::::||||||||||
Db    595 GTTGGGGAGCTCAACACCCGGCTCACAGATGATGTCTCCAAAATTAACGAAGGAATTGGT 654

Qy    118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
          |||||||||   :::   ||||||   |||:::||||||         |||      :::|||
Db    655 GACAAAATTGGAATGTTCTTTCAGGCAATGGCAACATTTTTTGGTGGTTTTATAATAGGA 714

Qy    138 LeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
                :::||||||||||||||||||||||||   |||:::   ||||||::::::
Db    715 TTTACTCGAGGCTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTG 774

Qy    158 SerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr 177
          ||||||            ::::::::::::   |||   |||   |||||||||   ||||||
Db    775 TCAGCTGGTATTTGGGCAAAGATATTATCTTCATTTACTGATAAAGAACTTCAGGCATAT 834

Qy    178 SerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197
          :::||||||||||||||||||||||||||||||||::::::||||||||||||||||||||||||
Db    835 GCAAAAGCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCATTT 894

Qy    198 ArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrp 217
                |||:::|||||||||:::
Db    895 GGAGGACAAAAGAAAGAACTTGAA------------------------------------ 918

Qy    218 PheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsn 237
                                                        ||||||      |||
Db    919 -------------------------------------------AGGTACAACAACAAT 933

Qy    238 LeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGly 257
          |||::::::||||||      ||||||||||:::   |||   :::   :::|||:::|||
Db    934 TTAGAAGAAGCCAAAAGACTTGGGATAAAGAAAGCGATCACAGCCAACATTTCCATGGGC 993

Qy    258 AlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeu 277
          |||   :::      :::      :::|||   ||||||||||||||||||||||||||||||
Db    994 GCAGCTTTCCTTCTTATCTATGCATCATATGCTCTGGCCTTCTGGTATGGAACGTCCTTG 1053

Qy    278 IleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerVal 297
          :::::::::   |||      |||:::||||||   ||||||   |||||||||:::|||
Db   1054 GTCATCTCAAAAGAA------TATTCTATTGGACAAGTGCTCACTGTCTTCTTTGCTGTA 1107

Qy    298 IleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAla 317
          :::      ::::::   ||||||   |||   |||:::      |||   ||||||   |||
Db   1108 TTAATTGGGGCATTCAGTATTGGACAGGCATCTCCAAATATCGAAGCCTTTGCCAATGCA 1167

Qy    318 ArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPhe 337
          ||||||||||||:::   ||||||   :::||||||   ||||||||||||||||:::|||
Db   1168 AGAGGAGCAGCCTATGAAATCTTCAACATAATTGATAATAAGCCCAGTATTGACAGCTTC 1227

Qy    338 SerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSer 357
          |||      ||||||||||||::::::|||:::|||   :::||||||||||||:::
Db   1228 TCAAAGAATGGGTACAAACCAGACAACATTAAAGGAAATTTGGAATTCAAAAATATTCAC 1287
```

```
Qy    358 PheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLys 377
Db   1288 TTCAGTTACCCATCTCGAAAAGACGTTCAGATCTTGAAGGGCCTCAACCTGAAGGTGCAG 1347

Qy    378 SerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGln 397
Db   1348 AGCGGACAGACAGTGGCCCTGGTTGGCAACAGTGGCTGTGGAAAAGCACCACTGTCCAG 1407

Qy    398 LeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIle 417
Db   1408 CTGCTGCAGAGGCTCTACGACCCCACAGAGGGCGTGGTCAGTATCGACGGTCAGGACATC 1467

Qy    418 ArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProVal 437
Db   1468 AGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTCGTGAGTCAGGAACCTGTG 1527

Qy    438 LeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGlu 457
Db   1528 TTGTTTGCCACCACTATCGCTGAAAACATTCGCTATGGCCGAGAAAATGTCACTATGGAC 1587

Qy    458 GluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsn 477
Db   1588 GAGATCGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTCATCATGAAACTGCCCCAT 1647

Qy    478 LysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArg 497
Db   1648 AAATTTGACACTCTTGTTGGTGAGAGAGGAGCACAGCTGAGTGGAGGACAGAAACAGAGA 1707

Qy    498 IleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThr 517
Db   1708 ATCGCCATTGCTCGTGCCCTGGTCCGCAACCCCAAGATCCTTTTGTTGGATGAGGCGACA 1767

Qy    518 SerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLys 537
Db   1768 TCAGCCTTGGACACAGAAAGTGAAGCCGTGGTTCAGGCTGCTCTGGATAAGGCTAGAGAA 1827

Qy    538 GlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIle 557
Db   1828 GGCCGGACTACCATTGTGATAGCCCACCGCTTGTCTACAGTTCGAAATGCTGACATCATA 1887

Qy    558 ValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLys 577
Db   1888 GCTGGGTTTGATGGTGGTGTCATTGTGGAGCAAGGAAATCATGAAGAGCTCATGAGAGAG 1947

Qy    578 ArgGlyLeuTyrTyrSerLeuValMetSerGln--------------------------- 588
Db   1948 AAGGGCATTTACTTCAAACTTGTCATGACACAGACAGCAGGAAATGAAATTGAATTAGGA 2007

Qy    589 ---AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 607
Db   2008 AATGAAGTTGGTGAGTCTAAAAATGAAATTGATAATTTAGACATGTCTTCAAAAGATTCA 2067

Qy    608 ThrAsnSerLeu------ProLeuHisSerValLysSerIleLysSerAspPheIleAsp 625
Db   2068 GCATCCAGTCTAATCAGAAGGAGATCAACTCGCAGAAGTATCCGTGGACCACATGACCAA 2127

Qy    626 LysAlaGluGluSerThr---GlnSerLysGluIleSerLeuProGluValSerLeuLeu 644
Db   2128 GACAGGAAGCTTAGTACCAAAGAGGCCTTGGATGAAGATGTACCTCCAATTTCCTTTTGG 2187

Qy    645 LysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 664
Db   2188 CGTATCCTAAAGTTGAATTCATCTGAATGGCCATATTTGTGGTTGGTATATTCTGTGCC 2247

Qy    665 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 684
Db   2248 ATAGTAAATGGAGCCTTGCAACCAGCATTCTCAATAATATTTTCTAAGGTTGTAGGGGTT 2307

Qy    685 Phe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 703
Db   2308 TTCACAAGAAATACTGATGATGAAACCAAACGACATGATAGCAACTTGTTTTCCTTATTG 2367

Qy    704 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 723
Db   2368 TTTCTGATCCTTGGGGTCATTTCTTTTATTACATTTTTTCTTCAGGGCTTCACATTTGGC 2427

Qy    724 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 743
Db   2428 AAAGCTGGAGAGATCCTCACCAAGCGACTCCGATACATGGTTTTCAAATCCATGCTGAGA 2487

Qy    744 GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 763
Db   2488 CAGGATGTGAGCTGGTTTGATAACCCTAAAAACACCACTGGAGCATTGACCACCAGGCTC 2547

Qy    764 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 783
Db   2548 GCCAACGATGCTGGTCAAGTTAAAGGGGCTACAGGAGCCAGACTTGCTGTCATTACCCAG 2607

Qy    784 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr 803
Db   2608 AATATAGCAAATCTTGGGACAGGAATCATCATATCCCTAATCTATGGCTGGCAGTTGACA 2667

Qy    804 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 823
Db   2668 CTTCTACTCTTAGCAATTGTCCCCATCATTGCAATAGCAGGAGTGGTTGAGATGAAAATG 2727

Qy    824 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 843
Db   2728 TTGTCTGGACAAGCACTAAAAGATAAGAAGGAGCTAGAAGGTTCTGGGAAGATTGCTACT 2787

Qy    844 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 863
Db   2788 GAAGCAATAGAGAACTTCCGCACTGTCGTCTCTTTGACTCGGGAGCAGAAGTTTGAAAAT 2847

Qy    864 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 883
Db   2848 ATGTATGCCCAGAGCTTGCAGATACCATACAGAAATGCTCTGAAGAAAGCACACGTCTTT 2907

Qy    884 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 903
Db   2908 GGGATCACATTCTCCTTCACCCAGGCCATGATGTATTTTTCGTATGCTGCTTGTTTCCGG 2967

Qy    904 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 923
Db   2968 TTTGGTGCTTACTTGGTGGCACGGGAATTAATGACATTTGAAAATGTTCTATTAGTATTC 3027

Qy    924 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 943
Db   3028 TCAGCTATTGTCTTTGGTGCCATGGCAGTTGGGCAGGTCAGTTCATTTGCTCCTGACTAT 3087

Qy    944 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 963
Db   3088 GCCAAAGCCAAAGTGTCGGCATCGCATATCATCATGATCATTGAAAAAGTCCCTTCCATT 3147

Qy    964 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 983
Db   3148 GACAGCTACAGCACCGGAGGCCTGAAGCCTAATACATTGGAAGGAAATGTGAAATTTAAT 3207

Qy    984 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 1003
Db   3208 GAAGTCGTGTTCAACTACCCCACCCGGCCTGACATCCCAGTGCTTCAGGGGCTGAACCTG 3267

Qy   1004 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 1023
Db   3268 GAGGTGAAGAAGGGCCAGACGCTGGCCCTGGTGGGCAGCAGTGGCTGCGGGAAGAGCACA 3327

Qy   1024 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 1043
Db   3328 GTGGTCCAGCTGCTCGAGCGGTTCTACGACCCCATGGCCGGAACAGTGTTTCTAGATGGC 3387

Qy   1044 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 1063
Db   3388 AAAGAAGTAAACCAACTAAATGTCCAGTGGCTCCGAGCACACCTAGGCATTGTGTCCCAA 3447

Qy   1064 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 1083
```

```
            |||||:::|||||:::||||||||||||||||||||||||||||||||||||||||||
Db   3448 GAGCCCATCCTGTTTGACTGCAGCATCGCTGAGAACATCGCTTATGGAGACAACAGCCGG 3507

Qy   1084 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 1103
          |||||      |||||||:::  ||||||      ||||||||||  ||||||
Db   3508 GTCGTGTCCCAGGATGAGATCGAGAGGGCGGCGAAGGAAGCCAACATCCACCAGTTCATT 3567

Qy   1104 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 1123
          |||  ||||||:::||||||||:::||||||  ||||||   |||||||||||
Db   3568 GAGTCACTGCCTGACAAATACAACACCAGAGTAGGAGACAAAGGGACTCAGCTGTCGGGT 3627

Qy   1124 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 1143
          ||||||||||||:::||||||||||||||||||:::::::::|||   ||||||||
Db   3628 GGGCAGAAACAACGCATCGCCATAGCACGCGCCCTCGTCAGACAGCCTCACATTTTGCTC 3687

Qy   1144 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 1163
          ||||||||||||||||||||||||   :::|||||||||||||||   ||||||
Db   3688 CTGGATGAAGCAACATCAGCCCTGGATACAGAGAGTGAAAAGGTTGTCCAGGAAGCTCTG 3747

Qy   1164 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 1183
          |||||||||||   ||||||||||:::|||:::   ||||||||||   ||||||
Db   3748 GACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATCGCTCACCGCCTGTCCACCATCCAG 3807

Qy   1184 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 1203
          ||||||||||||||||||||:::  ||||||||||:::||||||   |||||||||||
Db   3808 AACGCAGACTTGATCGTGGTGATTCAGAATGGCAAGGTCAAGGAGCATGGCACCCACCAG 3867

Qy   1204 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
          :::||||||   :::   ||||||||   :::|||:::   |||:::
Db   3868 CAGCTGCTGGCACAGAAAGGCATCTATTTCTCCATGGTCAGTGTGCAGGCT 3918

RESULT 7
CRUPGP1165
LOCUS       CRUPGP1165        4304 bp    mRNA    linear   ROD 27-APR-1993
DEFINITION  Chinese hamster p-glycoprotein mRNA (clone ADX165), complete cds.
ACCESSION   M59253
VERSION     M59253.1  GI:191154
KEYWORDS    multidrug resistance glycoprotein; p-glycoprotein.
SOURCE      Chinese hamster, cDNA to mRNA, clone ADX165.
  ORGANISM  Cricetulus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
            Cricetulus.
REFERENCE   1  (bases 1 to 4304)
  AUTHORS   Devine,S.E., Hussain,A., Davide,J.P. and Melera,P.W.
  TITLE     Full length and alternatively spliced pgp1 transcripts in
            multidrug-resistant Chinese hamster lung cells
  JOURNAL   J. Biol. Chem. 266 (7), 4545-4555 (1991)
  MEDLINE   91154265
   PUBMED   1671863
FEATURES             Location/Qualifiers
     source          1..4304
                     /organism="Cricetulus sp."
                     /db_xref="taxon:10031"
                     /cell_type="fibroblast"
                     /tissue_type="lung"
     gene            1..4304
                     /gene="pgp-1"
     CDS             111..3941
                     /gene="pgp-1"
                     /codon_start=1
                     /product="p-glycoprotein"
                     /protein_id="AAA37004.1"
                     /db_xref="GI:191155"
                     /translation="MEFEEDFSGRKDKNFLKMGRKSKKEKKEKKPVVSVFTMFRYAGW
                     LDRLYMLVGTLAAIIHGVALPLMMLVFGDMTDSFASVGNIPTNATNNATQVNASDIFG
                     KLEEEMTTYAYYYTGIGAGVLIVAYIQVSFWCLAAGRQIHKIRQKFFHAIMNQEIGWF
                     DVHDVGELNTRLTDDVSKINEGIGDKIGMFFQAMATFFGGFIIGFTRGWKLTLVILAI
                     SPVLGLSAGIWAKILSSFTDKELQAYAKAGAVAEEVLAAIRTVIAFGGQKKELERYNN
                     NLEEAKRLGIKKAITANISMGAAFLLIYASYALAFWYGTSLVISKEYSIGQVLTVFFA
                     VLIAPFSIGQASPNIEAFANARGAAYEIFNIIDNKPSIDSFSKNGYKPDNIKGNLEFK
                     NIHFSYPSRKDVQILKGLNLKVQSGQTVALVGNSGCGKSTTVQLLQRLYDPTEGVVSI
                     DGQDIRTINVRYLREIIGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKEANAYD
                     FIMKLPHKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTESEAVV
                     QAALDKAREGRTTIVIAHRLSTVRNADIIAGFDGGVIVEQGNHEELMREKGIYFKLVM
                     TQTAGNEIELGNEVGESKNEIDNLDMSSKDSASSLIRRRSTRRSIRGPHDQDRKLSTK
                     EALDEDVPPISFWRILKLNSSEWPYFVVGIFCAIVNGALQPAFSIIFSKVVGVFTRNT
                     DDETKRHDSNLFSLLFLILGVISFITFFLQGFTFGKAGEILTKRLRYMVFKSMLRQDV
                     SWFDNPKNTTGALTTRLANDAGQVKGATGARLAVITQNIANLGTGIIISLIYGWQLTL
                     LLAIVPIIAIAGVVEMKMLSGQALKDKKELEGSGKIATEAIENFRTVVSLTREQKFE
                     NMYAQSLQIPYRNALKKAHVFGITFSFTQAMMYFSYAACFRFGAYLVARELMTFENVL
                     LVFSAIVFGAMAVGQVSSFAPDYAKAKVSASHIIMIIEKVPSIDSYSTGGLKPNTLEG
                     NVKFNEVVFNYPTRPDIPVLQGLNLEVKKGQTLALVGSSGCGKSTVVQLLERFYDPMA
                     GTVFLDGKEVNQLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSQDEIERAA
                     KEANIHQFIESLPDKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEATSALD
                     TESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVIQNGKVKEHGTHQQLLAQKG
                     IYFSMVSVQAGAKR"
BASE COUNT     1275 a    895 c    1031 g    1103 t
ORIGIN      Map position 1q26.

Alignment Scores:
Pred. No.:               2.39e-228     Length:          4304
Score:                   3549.50       Matches:         679
Percent Similarity:      73.75%        Conservative:    248
Best Local Similarity:   54.02%        Mismatches:      259
Query Match:             57.63%        Indels:          71
DB:                      10            Gaps:            9

US-09-873-409-5 (1-1222) x CRUPGP1165 (1-4304)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          |||::::::|||   ||||||:::::::::::::|||      ||||||||||   |||
Db    258 ATGCTGGTGGGGACTCTCGCAGCTATTATCCATGGAGTTGCACTCCCGCTTATGATGCTG 317

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsn------ 38
          |||   |||:::|||:::|||:::      |||      :::   |||||
Db    318 GTGTTTGGAGACATGACAGACAGCTTTGCAAGTGTAGGAAACATCCCAACAAATGCTACT 377

Qy     38 ------------------------------------------------------------ 38

Db    378 AATAATGCCACTCAAGTCAATGCTTCAGACATCTTTGGCAAACTGGAGGAAGAAATGACC 437

Qy     39 ThrTyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIle 58
          ||||||::::::               ||||||   |||||||||        ||||||
Db    438 ACGTACGCCTAT---------------TACTACACAGGGATTGGTGCTGGTGTGCTTATA 482

Qy     59 PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIle 78
             |||||||||:::|||   |||   :::   |||   ||||||      :::|||
Db    483 GTTGCCTACATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATT 542

Qy     79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp 98
          |||::::::|||||||||:::::::::   |||:::|||||||||||||||      |||
Db    543 AGACAGAAGTTTTTTCATGCTATAATGAATCAGGAGATAGGCTGGTTTGACGTGCACGAC 602

Qy     99 IleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly 117
          :::||||||||||||||||||:::|||   |||:::   ||||||::::::|||||||||
Db    603 GTTGGGGAGCTCAACACCCGGCTCACAGATGATGTCTCCAAAATTAACGAAGGAATTGGT 662

Qy    118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
          |||||||||   :::   ||||||   |||:::||||||      |||      :::|||
Db    663 GACAAAATTGGAATGTTCTTTCAGGCAATGGCAACATTTTTTGGTGGTTTTATAATAGGA 722

Qy    138 LeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
             :::||||||||||||||||||||   |||:::   ||||||::::::
Db    723 TTTACTCGAGGCTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTG 782

Qy    158 SerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr 177
          ||||||         ::::::::::::   |||   |||   |||||||||   ||||||
Db    783 TCAGCTGGTATTTGGGCAAAGATATTATCTTCATTTACTGATAAAGAACTTCAGGCATAT 842

Qy    178 SerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197
          :::||||||||||||||||||||||||||||||::::::||||||||||||||||||||
Db    843 GCAAAAGCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCATTT 902

Qy    198 ArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrp 217
             |||:::|||||||||:::
```

```
Db    903 GGAGGACAAAAGAAAGAACTTGAA------------------------------------ 926
Qy    218 PheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsn 237
Db    927 ---------------------------------------------AGGTACAACAACAAT 941
Qy    238 LeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGly 257
Db    942 TTAGAAGAAGCCAAAAGACTTGGGATAAAGAAAGCGATCACAGCCAACATTTCCATGGGC 1001
Qy    258 AlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeu 277
Db   1002 GCAGCTTTCCTTCTTATCTATGCATCATATGCTCTGGCCTTCTGGTATGGAACGTCCTTG 1061
Qy    278 IleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerVal 297
Db   1062 GTCATCTCAAAAGAA------TATTCTATTGGACAAGTGCTCACTGTCTTCTTTGCTGTA 1115
Qy    298 IleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAla 317
Db   1116 TTAATTGCGCCATTCAGTATTGGACAGGCATCTCCAAATATCGAAGCCTTTGCCAATGCA 1175
Qy    318 ArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPhe 337
Db   1176 AGAGGAGCAGCCTATGAAATCTTCAACATAATTGATAATAAGCCCAGTATTGACAGCTTC 1235
Qy    338 SerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSer 357
Db   1236 TCAAAGAATGGGTACAAACCAGACAACATTAAAGGAAATTTGGAATTCAAAAATATTCAC 1295
Qy    358 PheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLys 377
Db   1296 TTCAGTTACCCATCTCGAAAAGACGTTCAGATCTTGAAGGGCCTCAACCTGAAGGTGCAG 1355
Qy    378 SerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGln 397
Db   1356 AGCGGACAGACAGTGGCCCTGGTTGGCAACAGTGGCTGTGGGAAAAGCACCACTGTCCAG 1415
Qy    398 LeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIle 417
Db   1416 CTGCTGCAGAGGCTCTACGACCCCACAGAGGGCGTGGTCAGTATCGACGGACAGGACATC 1475
Qy    418 ArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProVal 437
Db   1476 AGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTCGTGAGTCAGGAACCTGTG 1535
Qy    438 LeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGlu 457
Db   1536 TTGTTTGCCACCACTATCGCTGAAAACATTCGCTATGGCCGAGAAAATGTCACTATGGAC 1595
Qy    458 GluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsn 477
Db   1596 GAGATCGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTCATCATGAAACTGCCCCAT 1655
Qy    478 LysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArg 497
Db   1656 AAATTTGACACTCTTGTTGGTGAGAGAGGAGCACAGCTGAGTGGAGGACAGAAACAGAGA 1715
Qy    498 IleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThr 517
Db   1716 ATCGCCATTGCTCGTGCCCTGGTCCGCAACCCCAAGATCCTTTTGTTGGATGAGGCGACA 1775
Qy    518 SerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLys 537
Db   1776 TCAGCCTTGGACACAGAAAGTGAAGCCGTGGTTCAGGCTGCTCTGGATAAGGCTAGAGAA 1835
Qy    538 GlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIle 557
Db   1836 GGCCGGACTACCATTGTGATAGCCCACCGCTTGTCTACAGTTCGAAATGCTGACATCATA 1895
Qy    558 ValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLys 577
Db   1896 GCTGGGTTTGATGGTGGTGTCATTGTGGAGCAAGGAAATCATGAAGAGCTCATGAGAGAG 1955

Qy    578 ArgGlyLeuTyrTyrSerLeuValMetSerGln--------------------------- 588
Db   1956 AAGGGCATTTACTTCAAACTTGTCATGACACAGACAGCAGGAAATGAAATTGAATTAGGA 2015
Qy    589 ---AspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 607
Db   2016 AATGAAGTTGGTGAGTCTAAAAATGAAATTGATAATTTAGACATGTCTTCAAAAGATTCA 2075
Qy    608 ThrAsnSerLeu------ProLeuHisSerValLysSerIleLysSerAspPheIleAsp 625
Db   2076 GCATCCAGTCTAATCAGAAGGAGATCAACTCGCAGAAGTATCCGTGGACCACATGACCAA 2135
Qy    626 LysAlaGluGluSerThr---GlnSerLysGluIleSerLeuProGluValSerLeuLeu 644
Db   2136 GACAGGAAGCTTAGTACCAAAGAGGCCTTGGATGAAGATGTACCTCCAATTTCCTTTTGG 2195
Qy    645 LysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 664
Db   2196 CGTATCCTAAAGTTGAATTCATCTGAATGGCCATATTTTGTGGTTGGTATATTCTGTGCC 2255
Qy    665 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 684
Db   2256 ATAGTAAATGGAGCCTTGCAACCAGCATTCTCAATAATATTTTCTAAGGTTGTAGGGGTT 2315
Qy    685 Phe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIle 703
Db   2316 TTCACAAGAAATACTGATGATGAAACCAAACGACATGATAGCAACTTGTTTTCCTTATTG 2375
Qy    704 PheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGly 723
Db   2376 TTTCTGATCCTTGGGGTCATTTCTTTTATTACATTTTTCTTCAGGGCTTCACATTTGGC 2435
Qy    724 ArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyr 743
Db   2436 AAAGCTGGAGAGATCCTCACCAAGCGACTCCGATACATGGTTTTCAAATCCATGCTGAGA 2495
Qy    744 GlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeu 763
Db   2496 CAGGATGTGAGCTGGTTTGATAACCCTAAAAACACCACTGGAGCATTGACCACCAGGCTC 2555
Qy    764 AlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGln 783
Db   2556 GCCAACGATGCTGGTCAAGTTAAAGGGGCTACAGGAGCCAGACTTGCTGTCATTACCCAG 2615
Qy    784 AsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThr 803
Db   2616 AATATAGCAAATCTTGGGACAGGAATCATCATATCCCTAATCTATGGCTGGCAGTTGACA 2675
Qy    804 PheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAla 823
Db   2676 CTTCTACTCTTAGCAATTGTCCCCATCATTGCAATAGCAGGAGTGGTTGAGATGAAAATG 2735
Qy    824 MetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThr 843
Db   2736 TTGTCTGGACAAGCACTAAAAGATAAGAAGGAGCTAGAAGGTTCTGGGAAGATTGCTACT 2795
Qy    844 GluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGln 863
Db   2796 GAAGCAATAGAGAACTTCCGCACTGTCGTCTCTTTGACTCGGGAGCAGAAGTTTGAAAAT 2855
Qy    864 MetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIle 883
Db   2856 ATGTATGCCCAGAGCTTGCAGATACCATACAGAAATGCTCTGAAGAAAGCACACGTCTTT 2915
Qy    884 GlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArg 903
Db   2916 GGGATCACATTCTCCTTCACCCAGGCCATGATGTATTTTTCGTATGCTGCTTGTTTCCGG 2975
Qy    904 PheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPhe 923
Db   2976 TTTGGTGCTTACTTGGTGGCACGGGAATTAATGACATTTGAAAATGTTCTATTAGTATTC 3035
```

```
Qy     924 ThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyr 943
              :::||||||   :::|||||||||||:::|||:::                |||||:::|||
Db    3036 TCAGCTATTGTCTTTGGTGCCATGGCAGTTGGGCAGGTCAGTTCATTTGCTCCTGACTAT 3095

Qy     944 SerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIle 963
              :::|||||||||       |||:::|||:::        :::::||||||   |||:::|||
Db    3096 GCCAAAGCCAAAGTGTCGGCATCGCATATCATCATGATCATTGAAAAAGTCCCTTCCATT 3155

Qy     964 AspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArg 983
           ||||||   |||         |||   ||||||:::|||   |||||||||::::::|||
Db    3156 GACAGCTACAGCACCGGAGGCCTGAAGCCTAATACATTGGAAGGAAATGTGAAATTTAAT 3215

Qy     984 GluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeu 1003
           ||||||   |||   ||||||   |||||||||:::   :::|||:::||||||:::|||
Db    3216 GAAGTCGTGTTCAACTACCCCACCCGGCCTGACATCCCAGTGCTTCAGGGGCTGAACCTG 3275

Qy    1004 SerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThr 1023
            :::::::|||:::|||:::|||   ||||||||||||||||||||||||||||||||||||
Db    3276 GAGGTGAAGAAGGGCCAGACGCTGGCCCTGGTGGGCAGCAGTGGCTGCGGGAAGAGCACA 3335

Qy    1024 SerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 1043
           ||||||||||:::|||   |||||||||:::   |||   |||   ||||||
Db    3336 GTGGTCCAGCTGCTCGAGCGGTTCTACGACCCCATGGCCGGAACAGTGTTTCTAGATGGC 3395

Qy    1044 ValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGln 1063
              :::      :::||||||||||||||||||||:::   :::   ||||||   |||
Db    3396 AAAGAAGTAAACCAACTAAATGTCCAGTGGCTCCGAGCACACCTAGGCATTGTGTCCCAA 3455

Qy    1064 GluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArg 1083
           ||||||:::||||||:::||||||||||||||||||||||||||||||||||||||||||
Db    3456 GAGCCCATCCTGTTTGACTGCAGCATCGCTGAGAACATCGCTTATGGAGACAACAGCCGG 3515

Qy    1084 ValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIle 1103
           ||||||      |||||||||:::   ||||||      ||||||||||||   ||||||
Db    3516 GTCGTGTCCCAGGATGAGATCGAGAGGGCGGCGAAGGAAGCCAACATCCACCAGTTCATT 3575

Qy    1104 GluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGly 1123
           |||   ||||||:::|||||||||||||||:::||||||   ||||||   ||||||||||||
Db    3576 GAGTCACTGCCTGACAAATACAACACCAGAGTAGGAGACAAAGGGACTCAGCTGTCGGGT 3635

Qy    1124 GlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeu 1143
           ||||||||||||:::|||||||||||||||||||:::::::::|||   |||||||||
Db    3636 GGGCAGAAACAACGCATCGCCATAGCACGCGCCCTCGTCAGACAGCCTCACATTTTGCTC 3695

Qy    1144 LeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeu 1163
           ||||||||||||||||||||||||      :::||||||||||||||||   ||||||
Db    3696 CTGGATGAAGCAACATCAGCCCTGGATACAGAGAGTGAAAAGGTTGTCCAGGAAGCTCTG 3755

Qy    1164 AspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGln 1183
           ||||||||||||   |||||||||||:::|||:::   ||||||||||||   ||||||
Db    3756 GACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATCGCTCACCGCCTGTCCACCATCCAG 3815

Qy    1184 AsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGln 1203
           ||||||||||||||||||||:::      |||||||||:::|||||   ||||||||||||
Db    3816 AACGCAGACTTGATCGTGGTGATTCAGAATGGCAAGGTCAAGGAGCATGGCACCCACCAG 3875

Qy    1204 GluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
            :::||||||      :::   |||||||||   :::|||:::   |||:::
Db    3876 CAGCTGCTGGCACAGAAAGGCATCTATTTCTCCATGGTCAGTGTGCAGGCT 3926

RESULT 8
AX105082
LOCUS       AX105082                4279 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 26 from Patent WO0123540.
ACCESSION   AX105082
VERSION     AX105082.1  GI:13921232
KEYWORDS    .
SOURCE      dog.
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 4279)
  AUTHORS   Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and
            Patten,C.J.
  TITLE     P-glycoproteins and uses thereof
  JOURNAL   Patent: WO 0123540-A 26 05-APR-2001;
            GENTEST CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..4279
                     /organism="Canis familiaris"
                     /db_xref="taxon:9615"
     CDS             17..3862
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC37723.1"
                     /db_xref="GI:13921233"
                     /translation="MDPEGGRKGSAEKNFWKMGKKSKKKEKKEKKPTVSTFAMFRYSN
                     WLDRLYMLVGTMAAIIHGAALPLMMLVFGNMTDSFANAGISRNKTFPVIINESITNNT
                     QHFINHLEEEMTTYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQILKIRKQFFHAIMRQ
                     EIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSIATFFTGFIVGFTRGWKLTL
                     VILAISPVLGLSAAIWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFGGQKKEL
                     ERYNKNLEEAKGIGIKKAITANISIGAAFLLIYASYALAFWYGTSLVLSSEYTIGQVL
                     TVFFSVLIGAFSIGQASPSIEAFANARGAAYEIFKIIDNKPSIDSYSKSGHKPDNIKG
                     NLEFKNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQRLYDPTD
                     GMVCIDGQDIRTINVRHLREITGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKE
                     ANAYDFIMKLPNKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTE
                     SEAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELMKEKGIY
                     FKLVTMQTRGNEIELENATGESKSESDALEMSPKDSGSSLIKRRSTRRSIHAPQGQDR
                     KLGTKEDLNENVPPVSFWRILKLNSTEWPYFVVGIFCAIINGGLQPAFSIIFSRIIGI
                     FTRDEDPETKRQNSNMFSVLFLVLGIISFITFFLQGFTFGKAGEILTKRLRYMVFRSM
                     LRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAVITQNIANLGTGIIISLIYG
                     WQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRTVVSLTR
                     EQKFEYMYAQSLQVPYRNSLRKAHIFGVSFSITQAMMYFSYAGCFRFGAYLVANEFMN
                     FQDVLLVFSAIVFGAMAVGQVSSFAPDYAKAKVSAAHVIMIIEKSPLIDSYSPHGLKP
                     NTLEGNVTFNEVVFNYPTRPDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERF
                     YDPLAGSVLIDGKEIKHLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSHEE
                     IVQAAKEANIHHFIETLPEKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEA
                     TSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGKVKEHGTHQQL
                     LAQKGIYFSMVSVQAGAKR"
BASE COUNT     1296 a    833 c   1009 g   1141 t
ORIGIN

Alignment Scores:
Pred. No.:              2.8e-227        Length:         4279
Score:                  3533.50         Matches:        683
Percent Similarity:     73.13%          Conservative:   237
Best Local Similarity:  54.29%          Mismatches:     269
Query Match:            57.37%          Indels:         69
DB:                     6               Gaps:           9

US-09-873-409-5 (1-1222) x AX105082 (1-4279)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||::::::|||   :::|||:::::::::::|||||||||   ||||||||||||   |||
Db     167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
           |||   |||   |||:::|||:::      :::      :::
Db     227 GTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy      35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                                :::      ||||||      |||
Db     287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy      45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                          ||||||   |||||||||         |||:::
Db     347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy      60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
              |||||||||:::|||   |||   :::   |||   ||||||      :::||||||
Db     407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy      80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
           |||||||||||||:::::::::   |||:::|||||||||||||||         |||:::
Db     467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526
```

```
Qy    100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db    527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy    119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db    587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy    139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db    647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy    159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db    707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy    179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db    767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy    199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
Db    827 GGACAAAAGAAAGAACTTGAA--------------------------------------- 847

Qy    219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
Db    848 ------------------------------------------AGGTACAACAAAAATTTA 865

Qy    239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
Db    866 GAAGAAGCTAAAGGAATTGGGATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCC 925

Qy    259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
Db    926 GCTTTCTTATTGATCTATGCATCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTC 985

Qy    279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
Db    986 CTCTCCAGTGAA------TATACTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTA 1039

Qy    299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
Db   1040 ATTGGGGCTTTTAGTATTGGACAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGA 1099

Qy    319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
Db   1100 GGAGCAGCTTATGAAATCTTCAAGATAATTGACAATAAACCAAGCATTGACAGCTATTCG 1159

Qy    339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
Db   1160 AAGAGTGGACATAAACCAGATAATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTC 1219

Qy    359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
Db   1220 AGTTACCCTTCTCGAAAAGAAGTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGT 1279

Qy    379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
Db   1280 GGGCAGACAGTGGCGCTGGTTGGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTG 1339

Qy    399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
Db   1340 ATGCAGAGGCTCTATGACCCCACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGG 1399

Qy    419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db   1400 ACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTG 1459

Qy    439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
Db   1460 TTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAG 1519

Qy    459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
Db   1520 ATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTATCATGAAACTACCTAATAAA 1579

Qy    479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db   1580 TTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATC 1639

Qy    499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db   1640 GCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCA 1699

Qy    519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 538
Db   1700 GCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGC 1759

Qy    539 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558
Db   1760 CGGACTACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCCGATGTCATTGCT 1819

Qy    559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578
Db   1820 GGTTTTGATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAG 1879

Qy    579 GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 598
Db   1880 GGCATTTACTTCAAACTTGTCACAATGCAG---ACAAGAGGAAATGAAATTGAGTTAGAA 1936

Qy    599 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 618
Db   1937 AATGCCACTGGTGAATCCAAAAGTGAAAGTGATGCCTTGGAAATG---TCTCCAAAAGAT 1993

Qy    619 IleLysSerAspPheIle------------------------------------------ 624
Db   1994 TCAGGGTCCAGTTTAATAAAAAGAAGATCAACTCGCAGGAGTATACATGCACCACAAGGC 2053

Qy    625 ---AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643
Db   2054 CAAGACAGAAAGCTTGGTACAAAAGAGGACTTGAATGAGAATGTACCTCCAGTTTCCTTC 2113

Qy    644 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 663
Db   2114 TGGAGGATTCTGAAGCTGAACTCAACTGAATGGCCTTATTTTGTGGTTGGTATATTTTGT 2173

Qy    664 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 683
Db   2174 GCTATTATAAACGGAGGCCTGCAACCAGCATTTTCAATAATATTTTCAAGGATTATAGGG 2233

Qy    684 MetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 702
Db   2234 ATCTTTACCCGAGATGAGGATCCTGAAACAAAACGACAGAATAGTAACATGTTTTCTGTA 2293

Qy    703 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722
Db   2294 TTGTTTCTAGTCCTTGGAATTATTTCTTTTATTACATTTTTCCTCCAGGGCTTCACATTT 2353

Qy    723 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 742
Db   2354 GGCAAAGCTGGGGAGATCCTCACTAAGCGGCTTCGATACATGGTTTTCAGATCCATGCTG 2413

Qy    743 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIle 762
Db   2414 AGACAGGATGTCAGCTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACAACCAGG 2473

Qy    763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 782
Db   2474 CTTGCCAATGATGCGGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTCATTACC 2533

Qy    783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 802
Db   2534 CAGAATATAGCAAATCTTGGGACAGGCATTATTATATCCTTAATCTATGGTTGGCAATTA 2593

Qy    803 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 822
```

```
Db   2594 ACACTTTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAA 2653

Qy    823 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 842
            ::::::|||   |||   ||||||||:::||||||:::   ||||||||||||||
Db   2654 ATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCT 2713

Qy    843 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 862
          |||||||||:::||||||   ||||||:::||||||||||||||||||:::   ||||||
Db   2714 ACAGAAGCCATCGAAAACTTCCGAACTGTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAA 2773

Qy    863 GlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 882
             ||||||   :::   ||||||      :::||||||:::   :::||||||   |||
Db   2774 TACATGTATGCACAGAGTTTGCAAGTACCATACAGAAACTCTTTGAGGAAAGCACACATC 2833

Qy    883 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 902
             |||      ::::::   :::   |||   :::||||||:::||||||      |||
Db   2834 TTCGGGGTCTCATTTTCTATCACCCAGGCAATGATGTATTTTTCCTATGCTGGCTGTTTC 2893

Qy    903 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 922
          ||||||||||||||||||:::               |||      :::   :::   :::|||
Db   2894 CGGTTTGGTGCCTACTTGGTGGCAAATGAGTTCATGAACTTTCAGGATGTTCTTTTGGTA 2953

Qy    923 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 942
          |||:::||||||   :::||||||||||||:::|||:::            ||||||:::
Db   2954 TTCTCAGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAGGTCAGTTCATTTGCTCCTGAC 3013

Qy    943 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 962
          |||:::||||||||||||   ||||||||||:::      ::::::||||||   |||
Db   3014 TATGCCAAAGCCAAAGTATCAGCAGCCCACGTCATCATGATCATTGAAAAAAGCCCTCTG 3073

Qy    963 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 982
          ||||||||   |||      |||   ||||||:::|||   |||||||||:::   |||
Db   3074 ATTGACAGCTACAGCCCTCACGGCCTCAAGCCAAATACGTTGGAAGGAAATGTGACATTT 3133

Qy    983 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002
             ||||||   |||   ||||||   |||||||||:::   :::|||:::|||||||
Db   3134 AATGAGGTCGTGTTCAACTATCCCACTCGACCAGACATCCCCGTGCTCCAGGGGCTGAGC 3193

Qy   1003 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 1022
          |||   :::::::::|||:::|||:::|||   ||||||||||||||||||||||||
Db   3194 CTCGAGGTGAAGAAGGGCCAGACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGC 3253

Qy   1023 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1042
          |||   ||||||||||||:::|||   |||||||||:::   |||   ||||||   |||
Db   3254 ACAGTTGTTCAGCTCCTAGAGCGCTTCTATGACCCCTTGGCTGGTTCAGTGCTAATTGAT 3313

Qy   1043 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 1062
          |||   :::   |||   ||||||||||||||||||||||   :::   ||||||
Db   3314 GGCAAAGAGATAAAGCACCTGAATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCT 3373

Qy   1063 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082
          |||||||||:::||||||:::||||||||||||||||||||||||||||||||||||||||
Db   3374 CAGGAGCCCATCCTGTTTGACTGCAGCATTGCCGAGAACATTGCCTATGGAGACAACAGC 3433

Qy   1083 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1102
          |||||||||   :::||||||   :::||||||   ||||||||||||   |||
Db   3434 CGGGTCGTATCACATGAAGAGATTGTGCAGGCAGCCAAGGAGGCCAACATACACCACTTC 3493

Qy   1103 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122
          ||||||   ||||||||||||||||||||:::||||||   ||||||   |||||||||
Db   3494 ATCGAGACACTCCCTGAGAAATACAACACCAGAGTAGGAGACAAAGGAACCCAGCTCTCT 3553

Qy   1123 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 1142
          ||||||||||||||||||||:::||||||||||||||||:::::::::|||   ||||||
Db   3554 GGTGGCCAGAAACAGCGCATTGCCATAGCTCGCGCTCTTGTTAGACAGCCTCATATTTTG 3613

Qy   1143 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 1162
          ||||||||||||||||||||||||||||||   :::||||||||||||||||||   |||
Db   3614 CTTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCC 3673

Qy   1163 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 1182
          ||||||||||||||||   ||||||||||||:::|||:::   |||||||||||   |||
Db   3674 CTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATCGCCCACCGCTTGTCCACCATC 3733

Qy   1183 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 1202
          ||||||||||||||||||||||||   ||||||||:::||||||   |||||||||
Db   3734 CAGAATGCAGATTTAATAGTGGTGTTTCAGAATGGCAAAGTCAAGGAGCATGGCACACAT 3793

Qy   1203 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
          |||:::||||||      :::   |||||||||   :::|||:::   |||:::
Db   3794 CAACAGCTGCTGGCCCAGAAAGGCATCTATTTTTCCATGGTCAGTGTCCAGGCT 3847
```

RESULT 9
AX108656

```
LOCUS       AX108656               4195 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 3 from Patent WO0123565.
ACCESSION   AX108656
VERSION     AX108656.1  GI:13923888
KEYWORDS    .
SOURCE      crab-eating macaque.
  ORGANISM  Macaca fascicularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 4195)
  AUTHORS   Stocker,P.J., Steimel-Crespi,D.T. and Crespi,C.L.
  TITLE     P-glycoproteins from macaca fascicularis and uses thereof
  JOURNAL   Patent: WO 0123565-A 3 05-APR-2001;
            GENTEST CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..4195
                     /organism="Macaca fascicularis"
                     /db_xref="taxon:9541"
     CDS             100..3951
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC37765.1"
                     /db_xref="GI:13923889"
                     /translation="MDLEGDRNGGAEKKNFFKLNNKSKKDKKERKPTVSVFSMFRYSN
                     WLDKLYMVVGTLAAIIHGAGLPLMMLVFGDMTDTFANAGNLGDLGALLFNNTNSSNIT
                     DTVPVMNLEEDMTRYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQIHKIRKQFFHAIMR
                     QEIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSMATFFTGFIVGFTRGWKLT
                     LVILAISPVLGLSAAVWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFGGQKKE
                     LERYNKNLEEAKRIGIKKAITANISIGAAFLLIYASYALAFWYGTTLVLSKEYSIGQV
                     LTVFFSVLIGAFSVGQASPSIEAFANARGAAFEIFKIIDNKPSIDSYSKSGHKPDNIK
                     GNLEFRNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQRLYDPT
                     EGMVSVDGQDIRTINVRFLREIIGVVSQEPVLFATTIAENIRYGREDVTMDEIEKAVK
                     EANAYDFIMKLPQKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDT
                     ESEAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELMKEKGI
                     YFKLVTMQTAGNEIELENAADESKSEIDTLEMSSHDSGSSLIRKRSTRRSVRGSQGQD
                     RKLSTKEALDESIPPVSFWRIMKLNLTEWPYFVVGVFCAIINGGLQPAFAVIFSKIIG
                     IFTRNDDAETKRQNSNLFSLLFLVLGIVSFITFFLQGFTFGKAGEILTKRLRYMVFRS
                     MLRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAIITQNIANLGTGIIISLIY
                     GWQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRTVVSLT
                     QEQKFEHMYDQSLQVPYRNSLRKAHIFGITFSFTQAMMYFSYAGCFRFGAYLVAHSLM
                     SFEDVLLVFSAVVFGAMAVGQVSSFAPDYAKAKVSAAHIIMIIEKTPLIDSYSTEGLK
                     PNTLEGNVTFNEVVFNYPTRLDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLER
                     FYDPLAGKVLLDGKEIKQLNVQWLRAHLGIVSQEPILFDCSISENIAYGDNSRVVSQE
                     EIVRAAKEANIHAFIESLPNKYSTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDE
                     ATSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGRVKEHGTHQQ
                     LLAQKGIYFSMVSVQAGAKRQ"
BASE COUNT     1230 a    801 c   1039 g   1125 t
ORIGIN
```

```
Alignment Scores:
Pred. No.:               2.95e-227     Length:         4195
Score:                   3533.00       Matches:        680
Percent Similarity:      73.55%        Conservative:   246
Best Local Similarity:   54.01%        Mismatches:     263
Query Match:             57.36%        Indels:         70
DB:                      6             Gaps:           10
```

US-09-873-409-5 (1-1222) x AX108656 (1-4195)

```
Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          |||::::::|||   ||||||:::::::::::::||||||   ||||||||||||   |||
```

```
Db    250 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGAGCTGGACTTCCTCTCATGATGCTG 309
Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSer--------------------GlyCys 33
Db    310 GTGTTTGGAGACATGACGGATACCTTTGCAAATGCAGGAAATTTAGGAGATTTAGGAGCT 369
Qy     34 LeuValGlnThrAsnThrTyrSer----------------------------------- 41
Db    370 CTGTTGTTTAACAACACTAATAGCAGTAATATCACTGATACAGTGCCCGTCATGAATCTG 429
Qy     42 ---------PhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIle 58
Db    430 GAGGAAGATATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTT 489
Qy     59 PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIle 78
Db    490 GCTGCTTACATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATT 549
Qy     79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp 98
Db    550 AGAAAACAGTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGAT 609
Qy     99 IleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly 117
Db    610 GTTGGGGAGCTTAACACCCGGCTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGT 669
Qy    118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
Db    670 GACAAAATTGGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGA 729
Qy    138 LeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
Db    730 TTTACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTG 789
Qy    158 SerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr 177
Db    790 TCAGCTGCAGTCTGGGCAAAGATACTGTCTTCATTTACTGATAAAGAACTCTTAGCTTAT 849
Qy    178 SerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197
Db    850 GCAAAAGCTGGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTT 909
Qy    198 ArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrp 217
Db    910 GGAGGACAAAAGAAAGAACTCGAA------------------------------------ 933
Qy    218 PheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsn 237
Db    934 ---------------------------------------------AGGTACAACAAAAAT 948
Qy    238 LeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGly 257
Db    949 TTAGAAGAAGCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGT 1008
Qy    258 AlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeu 277
Db   1009 GCTGCTTTCCTGCTTATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTG 1068
Qy    278 IleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerVal 297
Db   1069 GTCCTCTCAAAGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTA 1122
Qy    298 IleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAla 317
Db   1123 TTAATTGGGGCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCA 1182
Qy    318 ArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPhe 337
Db   1183 AGAGGAGCAGCTTTTGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTAT 1242
Qy    338 SerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSer 357
Db   1243 TCGAAGAGTGGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCAC 1302

Qy    358 PheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLys 377
Db   1303 TTCAGTTACCCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAG 1362
Qy    378 SerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGln 397
Db   1363 AGTGGGCAGACGGTGGCCCTGGTTGGAAACAGCGGCTGTGGGAAGAGCACAACGGTCCAG 1422
Qy    398 LeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIle 417
Db   1423 CTGATGCAGAGGCTTTATGACCCCACAGAGGGCATGGTCAGTGTTGATGGACAGGATATT 1482
Qy    418 ArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProVal 437
Db   1483 AGGACCATAAACGTAAGGTTTCTACGGGAAATCATCGGTGTGGTGAGTCAGGAACCTGTA 1542
Qy    438 LeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGlu 457
Db   1543 TTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGTCGTGAAGATGTCACCATGGAT 1602
Qy    458 GluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsn 477
Db   1603 GAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCAG 1662
Qy    478 LysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArg 497
Db   1663 AAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGCTGAGTGGTGGGCAGAAGCAGAGG 1722
Qy    498 IleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThr 517
Db   1723 ATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGACGAGGCCACG 1782
Qy    518 SerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLys 537
Db   1783 TCAGCCTTGGACACAGAAAGTGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAA 1842
Qy    538 GlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIle 557
Db   1843 GGTCGGACCACCATTGTGATAGCTCATCGTTTGTCTACGGTTCGTAATGCCGACGTCATC 1902
Qy    558 ValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLys 577
Db   1903 GCTGGTTTCGATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAGCTCATGAAAGAG 1962
Qy    578 ArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIle--------------------- 590
Db   1963 AAAGGCATTTACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAATTGAATTAGAA 2022
Qy    591 LysLysAlaAspGlu---------GlnMetGluSerMetThrTyrSerThrGluArgLys 607
Db   2023 AATGCAGCTGATGAATCCAAAAGTGAAATTGATACCTTGGAAATGTCTTCACATGATTCA 2082
Qy    608 ThrAsnSerLeu---------------ProLeuHisSerValLysSerIleLysSerAsp 622
Db   2083 GGATCCAGTCTAATAAGAAAAAGATCCACTCGTAGGAGTGTCCGTGGATCACAAGGCCAA 2142
Qy    623 PheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSer 642
Db   2143 ------GACAGAAAGCTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCC 2196
Qy    643 LeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeu 662
Db   2197 TTTTGGAGGATTATGAAGCTAAATTTAACTGAGTGGCCTTATTTTGTTGTTGGTGTATTT 2256
Qy    663 AlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIle 682
Db   2257 TGTGCCATTATAAATGGAGGTCTGCAACCAGCATTTGCAGTAATATTTTCAAAGATTATA 2316
Qy    683 ThrMetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSer 701
Db   2317 GGGATTTTTACAAGAAATGATGATGCCGAAACAAAACGACAGAATAGTAACTTGTTTTCA 2376
```

```
Qy     702 MetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPhe 721
Db    2377 CTATTGTTTCTAGTCCTTGGAATTGTTTCTTTTATTACATTTTTCCTTCAGGGCTTCACA 2436

Qy     722 TyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMet 741
Db    2437 TTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATG 2496

Qy     742 LeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThr 761
Db    2497 CTCAGACAGGATGTGAGCTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACC 2556

Qy     762 IleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeu 781
Db    2557 AGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTATAATT 2616

Qy     782 ThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGlu 801
Db    2617 ACCCAGAATATAGCAAATCTTGGGACAGGAATAATTATATCCTTAATCTATGGTTGGCAA 2676

Qy     802 MetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThr 821
Db    2677 CTGACACTGTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATG 2736

Qy     822 AlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIle 841
Db    2737 AAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATC 2796

Qy     842 AlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPhe 861
Db    2797 GCTACTGAAGCAATAGAAAACTTCCGAACTGTTGTTTCTTTGACTCAGGAGCAGAAGTTT 2856

Qy     862 GluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGln 881
Db    2857 GAACATATGTATGATCAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACAC 2916

Qy     882 IleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGly 901
Db    2917 ATCTTTGGAATCACGTTTTCCTTCACGCAGGCAATGATGTATTTTTCCTATGCTGGATGT 2976

Qy     902 PheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIle 921
Db    2977 TTCCGGTTTGGAGCCTACTTGGTGGCACATAGTCTCATGAGCTTTGAGGATGTTCTGTTA 3036

Qy     922 ValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaPro 941
Db    3037 GTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCT 3096

Qy     942 GluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysPro 961
Db    3097 GACTATGCCAAAGCCAAAGTATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCT 3156

Qy     962 AsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGlu 981
Db    3157 TTGATTGACAGCTACAGCACAGAAGGCCTAAAGCCGAACACATTGGAAGGAAATGTCACA 3216

Qy     982 PheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeu 1001
Db    3217 TTTAATGAAGTTGTATTCAACTATCCCACCCGACTGGACATCCCAGTGCTTCAGGGGCTG 3276

Qy    1002 SerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLys 1021
Db    3277 AGCCTGGAAGTGAAGAAGGGCCAGACGCTGGCCCTGGTGGGCAGCAGTGGCTGTGGGAAG 3336

Qy    1022 SerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPhe 1041
Db    3337 AGCACGGTGGTCCAGCTCCTGGAGCGGTTCTATGACCCCTTGGCGGGGAAAGTGCTGCTT 3396

Qy    1042 AspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleVal 1061
Db    3397 GACGGCAAAGAAATAAAGCAACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTG 3456

Qy    1062 ProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsn 1081
Db    3457 TCCCAGGAGCCCATCCTGTTTGACTGCAGCATTAGTGAGAACATTGCCTATGGAGACAAC 3516

Qy    1082 SerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSer 1101
Db    3517 AGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCAGCCAAGGAGGCCAATATACACGCC 3576

Qy    1102 PheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeu 1121
Db    3577 TTCATCGAGTCACTGCCTAATAAATATAGCACCAGAGTAGGAGACAAAGGAACTCAGCTC 3636

Qy    1122 SerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIle 1141
Db    3637 TCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATT 3696

Qy    1142 LeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHis 1161
Db    3697 TTGCTTTTGGATGAAGCCACATCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAA 3756

Qy    1162 AlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAla 1181
Db    3757 GCCCTGGACAAAGCCAGAGAAGGCCGTACCTGCATTGTGATTGCTCACCGCCTGTCCACC 3816

Qy    1182 IleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThr 1201
Db    3817 ATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCACGGCACA 3876

Qy    1202 HisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db    3877 CATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3933

RESULT 10
AF257746
LOCUS       AF257746                4927 bp    mRNA    linear   ROD 09-MAY-2000
DEFINITION  Rattus norvegicus multidrug resistance protein 1a (Pgy1) mRNA,
            complete cds.
ACCESSION   AF257746
VERSION     AF257746.1  GI:7739772
KEYWORDS    .
SOURCE      Rattus norvegicus.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 4927)
  AUTHORS   Hooiveld,G.J.E.J., Wilms,J.W.J., Hagenbuch,B., Jansen,P.L.M.,
            Meijer,D.K.F. and Muller,M.
  TITLE     Cloning and functional characterization of the rat multidrug
            resistance protein Mdr1a
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4927)
  AUTHORS   Hooiveld,G.J.E.J., Wilms,J.W.J., Hagenbuch,B., Jansen,P.L.M.,
            Meijer,D.K.F. and Muller,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-APR-2000) Div. Gastroenterology & Hepatology,
            University Hospital Groningen, Hanzeplein 1, Groningen NL-9713 GZ,
            The Netherlands
FEATURES             Location/Qualifiers
     source          1..4927
                     /organism="Rattus norvegicus"
                     /strain="Wistar"
                     /db_xref="taxon:10116"
                     /sex="male"
                     /cell_type="hepatocyte"
                     /tissue_type="liver"
                     /dev_stage="adult"
     gene            1..4927
                     /gene="Pgy1"
     CDS             86..3904
                     /gene="Pgy1"
                     /note="ATP-binding cassette transport protein; Abcb1a"
                     /codon_start=1
                     /product="multidrug resistance protein 1a"
```

```
                     /protein_id="AAF69007.1"
                     /db_xref="GI:7739773"
                     /translation="MELEEDLNGRADKNFSKMGKKSKKEKKEKKPAVSVLTMFRYAGW
                     LDRFYMLLGTLAAIIHGIALPLMMLVFGDMTDSFANVGNNRSMSFYNATDIYAKLEDE
                     MTTYAYYYTGIGAGVLIVAYIQVSLWCLAAGRQIHKIRQKFFHAIMNQEIGWFDVHDV
                     GELNTRLTDDVSKINEGIGDKIGMFFQAMATFFGGFIIGFTRGWKLTLVILAISPVLG
                     LSAGIWAKILSSFTDKELQAYAKAGAVAEEVLAAIRTVIAFGGQKKELERYNNNLEEA
                     KRLGIKKAITANISMGAAFLLIYASYALAFWYGTSLVISKEYTIGQVLTVFFSVLIGA
                     FSVGQASPNIEAFANARGAAYEVFSIIDNKPSIDSFSKSGHKPDNIQGNLEFKNIHFS
                     YPSRKDVQILKGLNLKVKSGQTVALVGNSGCGKSTTVQLLQRLYDPIEGEVSIDGQDI
                     RTINVRYLREIIGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKEANAYDFIMKL
                     PHKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTESEAVVQAALD
                     KAREGRTTIVIAHRLSTVRNADVIAGFDGGVIVEQGNHDELMREKGIYFKLVMTQTAG
                     NEIELGNEACESKDGIDNVDMSSKDSGSSLIRRRSTRKSIRGPHDQDGELSTKEALDD
                     DVPPASFWRILKLNSTEWPYFVVGVFCAIINGGLQPAFSIIFSKVVGVFTKNDTPEIQ
                     RQNSNLFSLLFLILGIISFITFFLQGFTFGKAGEILTKRLRYMVFKSMLRQDISWFDD
                     PKNTTGALTTRLANDAAQVKGATGSRLAVITQNIANLGTGIIISLIYGWQLTLLLAI
                     VPIIAIAGVVEMKMLSGQALKDKKELEGSGKIATEAIENFRTVVSLTREQKFETMYAQ
                     SLQIPYRNALKKAHVFGITFSFTQAMMYFSYAACFRFGAYLVARELMTFENVLLVFSA
                     IVFGAMAVGQVSSFAPDYAKAKVSASHIIRIIEKIPEIDSYSTEGLKPNMLEGNVKFN
                     GVMFNYPTRPNIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERFYDPMAGTVFL
                     DGKEIKQLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSHEEIVKAAKEANI
                     HQFIDSLPEKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEATSALDTESEK
                     VVQEALDKAREGRTCIVIAHRLSTIQNADLIVVIQNGQVKEHGTHQQLLAQKGIYFSM
                     VSVQAGAKRS"
BASE COUNT     1474 a   1037 c   1190 g   1226 t
ORIGIN

Alignment Scores:
Pred. No.:              4.24e-227     Length:          4927
Score:                  3532.00       Matches:         682
Percent Similarity:     74.14%        Conservative:    244
Best Local Similarity:  54.60%        Mismatches:      263
Query Match:            57.35%        Indels:          60
DB:                     10            Gaps:            10

US-09-873-409-5 (1-1222) x AF257746 (1-4927)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db     233 ATGCTGCTGGGAACTCTGGCGGCCATTATCCATGGAATTGCGCTCCCACTTATGATGCTG 292

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db     293 GTCTTTGGAGACATGACAGATAGCTTTGCAAAT------GTAGGAAACAACCGTAGTATG 346

Qy      41 SerPhePheArgLeuThr------------------------------------------ 46
Db     347 AGTTTCTACAATGCTACAGACATATATGCCAAGCTGGAGGACGAAATGACCACGTACGCC 406

Qy      47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
Db     407 TACTATTACACGGGCATTGGTGCCGGTGTGCTCATCGTTGCCTACATCCAGGTTTCACTT 466

Qy      67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
Db     467 TGGTGCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAAGTTTTTCCATGCCATC 526

Qy      87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
Db     527 ATGAATCAGGAGATAGGCTGGTTTGACGTGCATGACGTTGGGGAGCTCAACACCCGGCTC 586

Qy     107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
Db     587 ACAGATGACGTCTCCAAAATTAATGAAGGAATTGGTGACAAAATTGGAATGTTCTTTCAG 646

Qy     126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
Db     647 GCAATGGCAACATTTTTTGGTGGTTTTATAATAGGATTTACTCGCGGCTGGAAGCTAACT 706

Qy     146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
Db     707 CTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTGTCAGCTGGTATTTGGGCAAAGATA 766

Qy     166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
Db     767 TTGTCTTCATTTACTGATAAGGAACTCCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAA 826

Qy     186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
Db     827 GAAGTCTTAGCAGCCATCAGAACTGTGATTGCCTTTGGAGGACAAAAGAAGGAACTTGAA 886

Qy     206 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeu 225
Db     886 ------------------------------------------------------------ 886

Qy     226 SerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 245
Db     887 ---------------------AGGTACAATAACAATTTGGAAGAAGCTAAAAGGCTTGGG 925

Qy     246 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 265
Db     926 ATAAAGAAAGCTATCACGGCCAACATTTCCATGGGTGCAGCTTTTCTGCTTATCTATGCA 985

Qy     266 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 285
Db     986 TCATATGCTCTGGCATTCTGGTATGGGACTTCCTTGGTCATCTCAAAAGAA------TAC 1039

Qy     286 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 305
Db    1040 ACTATTGGACAAGTGCTCACTGTCTTTTTTTCTGTATTAATTGGAGCATTCAGTGTTGGG 1099

Qy     306 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 325
Db    1100 CAGGCATCTCCAAATATTGAAGCCTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTC 1159

Qy     326 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 345
Db    1160 AGTATAATTGATAATAAGCCCAGTATAGACAGCTTCTCAAAGAGTGGGCACAAACCCGAC 1219

Qy     346 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 365
Db    1220 AACATACAAGGAAATTTGGAATTCAAAAATATTCACTTCAGTTACCCGTCTCGAAAAGAC 1279

Qy     366 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 385
Db    1280 GTTCAGATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTGGTT 1339

Qy     386 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 405
Db    1340 GGCAACAGTGGCTGTGGAAAAGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCC 1399

Qy     406 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 425
Db    1400 ATAGAGGGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTG 1459

Qy     426 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 445
Db    1460 CGGGAAATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCCACCACAATTGCCGAA 1519

Qy     446 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 465
Db    1520 AACATTCGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAA 1579

Qy     466 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 485
Db    1580 GCCAATGCCTATGATTTCATCATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAG 1639

Qy     486 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 505
Db    1640 AGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTC 1699

Qy     506 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 525
Db    1700 CGCAACCCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAA 1759
```

```
Qy    526 SerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 545
Db   1760 GCCGTGGTTCAGGCCGCTCTGGATAAGGCTAGAGAAGGCCGGACCACCATTGTGATAGCT 1819

Qy    546 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 565
Db   1820 CACCGCTTGTCTACAGTTCGCAATGCTGACGTCATTGCTGGTTTGATGGTGGTGTCATT 1879

Qy    566 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 585
Db   1880 GTGGAGCAAGGAAATCATGATGAGCTCATGAGAGAGAAAGGAATTTACTTCAAACTTGTC 1939

Qy    586 MetSerGln------------AspIleLysLysAlaAspGluGlnMetGluSer------ 599
Db   1940 ATGACTCAGACAGCAGGAAATGAAATTGAATTAGGAAATGAAGCTTGTGAATCTAAAGAT 1999

Qy    600 ------------MetThrTyrSerThrGluArgLysThrAsnSerLeu------ProLeu 613
Db   2000 GGAATTGATAATGTGGACATGTCTTCAAAAGATTCAGGATCCAGTCTAATAAGAAGAAGA 2059

Qy    614 HisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---Gln 632
Db   2060 TCAACTCGCAAAAGCATCCGTGGGCCACATGATCAAGACGGGGAACTTAGCACCAAAGAG 2119

Qy    633 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 652
Db   2120 GCTCTGGATGACGACGTACCTCCAGCTTCCTTTTGGCGGATCCTGAAGTTGAATTCAACT 2179

Qy    653 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 672
Db   2180 GAATGGCCTTATTTGTGGTTGGTGTATTTTGTGCCATAATAAATGGAGGCTTGCAACCA 2239

Qy    673 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 692
Db   2240 GCATTCTCCATAATATTTTCAAAGGTTGTAGGGGTTTTTACAAAAAATGACACCCCTGAA 2299

Qy    693 Leu---LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCys 711
Db   2300 ATCCAGCGGCAGAACAGCAACTTGTTTTCTTTATTGTTCCTGATCCTTGGGATCATCTCT 2359

Qy    712 PheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMet 731
Db   2360 TTCATTACGTTTTTCCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAG 2419

Qy    732 ArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGlu 751
Db   2420 CGACTCCGATACATGGTCTTCAAATCCATGCTGAGACAGGACATAAGCTGGTTTGATGAC 2479

Qy    752 LysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGln 771
Db   2480 CCTAAAAACACCACAGGAGCGCTGACCACCAGGCTTGCCAATGACGCTGCTCAAGTGAAA 2539

Qy    772 GlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSer 791
Db   2540 GGGGCTACAGGGTCTAGGCTTGCTGTTATTACCCAGAACATAGCAAATCTTGGGACAGGC 2599

Qy    792 ValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaPro 811
Db   2600 ATCATCATATCCCTGATCTACGGCTGGCAATTGACACTTTTACTCCTAGCAATTGTTCCC 2659

Qy    812 ValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAsp 831
Db   2660 ATCATTGCTATAGCAGGAGTGGTTGAAATGAAAATGTTGTCTGGACAAGCGCTGAAAGAT 2719

Qy    832 LysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThr 851
Db   2720 AAGAAGGAACTAGAAGGTTCTGGGAAGATCGCTACAGAAGCAATTGAAAACTTTCGCACT 2779

Qy    852 IleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThr 871
Db   2780 GTCGTCTCTTTGACTCGGGAGCAGAAGTTTGAAACTATGTATGCCCAGAGCTTGCAGATA 2839

Qy    872 GlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHis 891
Db   2840 CCATACAGAAATGCTTTGAAGAAAGCGCACGTCTTTGGGATCACTTTCTCCTTCACCCAG 2899

Qy    892 AlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAla 911
Db   2900 GCCATGATGTATTTCTCCTATGCTGCTTGTTTCCGGTTTGGTGCCTACTTGGTGGCACGA 2959

Qy    912 GlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMet 931
Db   2960 GAACTCATGACATTTGAAAATGTTCTGTTAGTATTCTCAGCTATTGTCTTTGGTGCCATG 3019

Qy    932 AlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAla 951
Db   3020 GCAGTGGGGCAGGTCAGTTCATTCGCTCCTGACTACGCGAAAGCCAAAGTCTCGGCATCC 3079

Qy    952 HisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLys 971
Db   3080 CACATCATCAGGATCATTGAGAAAATCCCTGAGATTGACAGCTACAGCACGGAGGGCTTG 3139

Qy    972 LysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCys 991
Db   3140 AAGCCTAATATGTTGGAAGGAAATGTGAAATTTAATGGAGTCATGTTCAACTATCCCACC 3199

Qy    992 ArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrVal 1011
Db   3200 CGACCCAACATCCCAGTGCTTCAGGGGCTGAGCCTAGAGGTGAAGAAAGGGCAGACGCTG 3259

Qy   1012 AlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeu 1031
Db   3260 GCCCTCGTGGGCAGCAGTGGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTTGAGCGCTTC 3319

Qy   1032 TyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnVal 1051
Db   3320 TATGACCCCATGGCCGGAACAGTGTTTCTAGATGGCAAAGAAATAAAGCAACTCAATGTC 3379

Qy   1052 GlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSer 1071
Db   3380 CAGTGGCTCCGCGCCCACCTGGGCATTGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGC 3439

Qy   1072 IleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLys 1091
Db   3440 ATCGCCGAGAACATTGCCTACGGAGACAACAGCCGTGTCGTGTCTCATGAGGAGATCGTG 3499

Qy   1092 GluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsn 1111
Db   3500 AAGGCAGCCAAGGAGGCCAACATCCACCAGTTCATCGACTCACTGCCTGAGAAATACAAC 3559

Qy   1112 ThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIle 1131
Db   3560 ACCAGAGTGGGAGACAAAGGGACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATCGCCATC 3619

Qy   1132 AlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeu 1151
Db   3620 GCGCGCGCCCTCGTCAGACAGCCTCACATCTTACTTCTGGATGAAGCGACATCAGCTCTG 3679

Qy   1152 AspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThr 1171
Db   3680 GATACGGAGAGTGAAAAGGTCGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACC 3739

Qy   1172 CysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeu 1191
Db   3740 TGCATTGTGATCGCGCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATT 3799

Qy   1192 HisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIle 1211
Db   3800 CAGAACGGCCAGGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCCCAGAAAGGCATC 3859

Qy   1212 TyrPheLysLeuValAsnAlaGlnSer 1220
Db   3860 TATTTCTCGATGGTCAGTGTGCAGGCT 3886
```

RESULT 11
AX322787

```
LOCUS       AX322787                3860 bp    DNA     linear   PAT 07-JAN-2002
DEFINITION  Sequence 1 from Patent WO0192877.
ACCESSION   AX322787
VERSION     AX322787.1  GI:18093766
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Sorrentino,B. and Schuetz,J.
  TITLE     Method of identifying and/or isolating stem cells
  JOURNAL   Patent: WO 0192877-A 1 06-DEC-2001;
            ST. JUDE CHILDREN'S RESEARCH HOSPITAL (US)
FEATURES             Location/Qualifiers
     source          1..3860
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1135 a    746 c    957 g   1022 t
ORIGIN

Alignment Scores:
Pred. No.:              3.9e-227       Length:          3860
Score:                  3530.50        Matches:         682
Percent Similarity:     74.00%         Conservative:    246
Best Local Similarity:  54.39%         Mismatches:      263
Query Match:            57.32%         Indels:          63
DB:                     6              Gaps:            10

US-09-873-409-5 (1-1222) x AX322787 (1-3860)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db      151 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210

Qy       21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
Db      211 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 270

Qy       32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db      271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330

Qy       44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db      331 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 390

Qy       62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db      391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450

Qy       82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db      451 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 510

Qy      102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db      511 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGGTATTGGTGACAAAATT 570

Qy      121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db      571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 630

Qy      141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db      631 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

Qy      161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db      691 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy      181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db      751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810

Qy      201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db      811 AAGAAAGAACTTGAA--------------------------------------------- 825

Qy      221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db      826 ------------------------------------AGGTACAACAAAAATTTAGAAGAA 849

Qy      241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db      850 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 909

Qy      261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db      910 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 969

Qy      281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db      970 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1023

Qy      301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db     1024 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1083

Qy      321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db     1084 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1143

Qy      341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db     1144 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1203

Qy      361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db     1204 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1263

Qy      381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db     1264 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1323

Qy      401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db     1324 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1383

Qy      421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db     1384 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1443

Qy      441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db     1444 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1503

Qy      461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db     1504 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1563

Qy      481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db     1564 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 1623

Qy      501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db     1624 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 1683

Qy      521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db     1684 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACC 1743

Qy      541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db     1744 ACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC 1803
```

```
Qy     561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
              |||||::::::  ||||||||  |||  ||||||||   ::::::|||:::
Db    1804 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATT 1863

Qy     581 TyrTyrSerLeuValMetSerGln-----------------AspIleLysLysAlaAsp 594
           |||:::  |||||  |||                        :::::::: |||
Db    1864 TACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 1923

Qy     595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
           ::::::   :::|||          :::   |||     :::|||
Db    1924 GATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGT 1983

Qy     615 SerVal-----------------LysSerIleLys---SerAspPheIleAspLysAla 627
           :::                     :::|||:::::: |||       |||:::
Db    1984 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAG 2043

Qy     628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
              :::  ::::::  :::  |||:::|||  |||||     :::|||:::
Db    2044 CTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATG 2103

Qy     648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
           ||||||||     |||||||||:::  |||:::|||       ::::::::|||
Db    2104 AAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAAT 2163

Qy     668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
           |||   :::   |||   |||:::|||||||||:::||||||||   :::|||
Db    2164 GGAGGCCTGCAACCAGCATTTGCAATAATATTTCAAAGATTATAGGGGTTTTACAAGA 2223

Qy     688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
              :::|||    |||   :::   ::::::   ::::::|||::::::|||:::
Db    2224 ATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCC 2283

Qy     707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
           |||||:::|||   |||::::::::|||::::::|||         :::|||:::|||
Db    2284 CTTGGAATTATTTCTTTTATTACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGA 2343

Qy     727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
           ||||||||||||   ||||||||::::::   |||::::::|||||   ||||||:::
Db    2344 GAGATCCTCACCAAGCGGCTCCGATACATGGTTTCCGATCCATGCTCAGACAGGATGTG 2403

Qy     747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
           :::|||||||||:::   :::|||:::|||||   |||||||||   ||||||    |||
Db    2404 AGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGAT 2463

Qy     767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
              ||||||::::::||||||   |||||||||:::   |||:::|||||||||
Db    2464 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCA 2523

Qy     787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
           |||:::|||        :::||||||||||||||||||||||||::::::|||   |||:::
Db    2524 AATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTC 2583

Qy     807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
           |||:::|||  |||::::::|||:::   |||::::::|||        ::::::|||
Db    2584 TTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGA 2643

Qy     827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
              |||   ||||||||:::||||||:::   ||||||||||||||||||||||:::
Db    2644 CAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 2703

Qy     847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
           ||||||   ||||||:::||||||||||||:::|||:::   ||||||   ||||||
Db    2704 GAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCT 2763

Qy     867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
           :::   ||||||      :::|||||||:::   :::|||||||   |||   |||
Db    2764 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 2823

Qy     887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
           ::::::|||:::   |||   :::||||||:::||||||      |||||||||||||||
Db    2824 TTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCC 2883

Qy     907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
           ||||||:::            |||:::   |||   :::   :::||||||:::|||:::
Db    2884 TACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTT 2943

Qy     927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
           :::||||||||||||:::|||:::              ||||||:::|||:::||||||
Db    2944 GTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCC 3003

Qy     947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
           |||      ||||||||||:::      ::::::||||||   |||   |||||||||
Db    3004 AAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTAC 3063

Qy     967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
           |||   ||||||      |||:::|||   |||||||||:::      |||   ||||||
Db    3064 AGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTA 3123

Qy     987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
           |||   ||||||   |||||||||:::   :::|||:::||||||||||||     ::::::
Db    3124 TTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAG 3183

Qy    1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
           :::|||:::|||:::|||   ||||||||||||||||||||||||||||   ||||||
Db    3184 AAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAG 3243

Qy    1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
           ||||||:::|||   |||||||||:::   |||:::|||||||   ||||||   :::
Db    3244 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATA 3303

Qy    1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
           |||   |||||||||||||||||||||:::   :::   ||||||   |||||||||:::
Db    3304 AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATC 3363

Qy    1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
           ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db    3364 CTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCA 3423

Qy    1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
              :::||||||      ||||||      ||||||||||||:::|||||||||   |||
Db    3424 CAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTG 3483

Qy    1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
           |||   |||||:::|||:::||||||   ||||||   ||||||||||||||||||
Db    3484 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAA 3543

Qy    1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
           ||||||:::||||||||||||||||||:::::::::|||   ||||||||||||||||||
Db    3544 CAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAA 3603

Qy    1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
           ||||||||||||||||||   :::||||||||||||||||||   ||||||||||||
Db    3604 GCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCC 3663

Qy    1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
           |||   ||||||||||||:::|||:::   ||||||||||||   ||||||||||||   |||
Db    3664 AGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGAC 3723

Qy    1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
           ||||||||||||      ||||||:::::::||||||   ||||||||||||:::||||||
Db    3724 TTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTG 3783

Qy    1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
              :::   |||||||||   :::|||:::   |||:::
Db    3784 GCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3825
```

```
RESULT 12
AX105078
LOCUS       AX105078                4279 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 22 from Patent WO0123540.
ACCESSION   AX105078
VERSION     AX105078.1  GI:13921228
KEYWORDS    .
SOURCE      dog.
  ORGANISM  Canis familiaris
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 4279)
  AUTHORS   Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and
            Patten,C.J.
  TITLE     P-glycoproteins and uses thereof
  JOURNAL   Patent: WO 0123540-A 22 05-APR-2001;
            GENTEST CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..4279
                     /organism="Canis familiaris"
                     /db_xref="taxon:9615"
     CDS             17..3862
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC37721.1"
                     /db_xref="GI:13921229"
                     /translation="MDPEGGRKGSAEKNFWKMGKKSKKNEKKEKKPTVSTFAMFRYSN
                     WLDRLYMLVGTMAAIIHGAALPLMMLVFGNMTDSFANAGISRNKTFPVIINESITNNT
                     QHFINHLEEEMTTYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQILKIRKQFFHAIMRQ
                     EIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSIATFFTGFIVGFTRGWKLTL
                     VILAISPVLGLSAAIWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFGGQKKEL
                     ERYNKNLEEAKGIGIKKAITANISIGAAFLLIYASYALAFWYGTSLVLSSEYSIGQVL
                     TVFFSVLIGAFSIGQASPSIEAFANARGAAYEIFKIIDNKPSIDSYSKSGHKPDNIKG
                     NLEFKNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQRLYDPTD
                     GMVCIDGQDIRTINVRHLREITGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKE
                     ANAYDFIMKLPNKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTE
                     SEAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELMKEKGIY
                     FKLVTMQTRGNEIELENATGESKSESDALEMSPKDSGSSLIKRRSTRRSIHAPQGQDR
                     KLGTKEDLNENVPPVSFWRILKLNSTEWPYFVVGIFCAIINGGLQPAFSIIFSRIIGI
                     FTRDEDPETKRQNSNMFSVLFLVLGIISFITFFLQGFTFGKAGEILTKRLRYMVFRSM
                     LRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAVITQNIANLGTGIIISLIYG
                     WQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRTVVSLTR
                     EQKFEYMYAQSLQVPYRNSLRKAHIFGVSFSITQAMMYFSYAGCFRFGAYLVANEFMN
                     FQDVLLVFSAIVFGAMAVGQVSSFAPDYAKAKVSAAHVIMIIEKSPLIDSYSPHGLKP
                     NTLEGNVTFNEVVFNYPTRPDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERF
                     YDPLAGSVLIDGKEIKHLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSHEE
                     IMQAAKEANIHHFIETLPEKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEA
                     TSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGKVKEHGTHQQL
                     LAQKGIYFSMVSVQAGAKR"
BASE COUNT     1295 a    833 c   1008 g   1143 t
ORIGIN

Alignment Scores:
Pred. No.:              4.45e-227       Length:         4279
Score:                  3530.50         Matches:        682
Percent Similarity:     73.13%          Conservative:   238
Best Local Similarity:  54.21%          Mismatches:     269
Query Match:            57.32%          Indels:         69
DB:                     6               Gaps:           9

US-09-873-409-5 (1-1222) x AX105078 (1-4279)

Qy          1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
              |||::::::||| :::|||:::::::::::||||||  |||||||||||   |||
Db        167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy         21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
              |||   |||   |||:::|||:::     :::     :::
Db        227 GTTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy         35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                                   :::      |||||     |||
Db        287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy         45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                             ||||||   |||||||||       |||:::
Db        347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy         60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
                 |||||||||:::|||  |||    :::   |||   ||||||     :::||||||
Db        407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy         80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
              ||||||||||||||:::::::::  |||:::||||||||||||||         |||:::
Db        467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy        100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
              |||||||||||||||||:::|||   |||:::   ||||||:::::||||||||||||
Db        527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy        119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
              ||||||   :::   ||||||:::::::::||||||      |||      ||||||
Db        587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy        139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
                 :::||||||||||||||||||||||   |||:::   ||||||::::::     |||
Db        647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy        159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
              ||||||      ::::::::::::   |||   |||   |||||||||   ||||||:::
Db        707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy        179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
              |||||||||||||||||||||||||||||||::::::|||||||||||||||||||||
Db        767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy        199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
                 |||:::|||||||||:::
Db        827 GGACAAAAGAAAGAACTTGAA--------------------------------------- 847

Qy        219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
                                                   ||||||    :::||||||
Db        848 ---------------------------------------AGGTACAACAAAAATTTA 865

Qy        239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
              ::::::||||||      |||||||||:::   |||    :::   :::|||:::||||||
Db        866 GAAGAAGCTAAAGGAATTGGGATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCC 925

Qy        259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
                 :::      :::   :::|||   |||||||||||||||||||||||||||:::
Db        926 GCTTTCTTATTGATCTATGCATCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTC 985

Qy        279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
              |||:::   |||      |||:::||||||   ||||||   ||||||||||||||:::
Db        986 CTCTCCAGTGAA------TATTCTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTA 1039

Qy        299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
                 ::::::   ||||||   |||   |||      |||   ||||||   ||||||
Db       1040 ATTGGGGCTTTTAGTATTGGACAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGA 1099

Qy        319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
              |||||||||:::   ||||||::::::||||||   ||||||||||||||::::::|||
Db       1100 GGAGCAGCTTATGAAATCTTCAAGATAATTGACAATAAACCAAGCATTGACAGCTATTCG 1159

Qy        339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
                 :::|||:::||||||||||:::::::|||:::|||   :::|||||||||||||||   |||
Db       1160 AAGAGTGGACATAAACCAGATAATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTC 1219

Qy        359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
              :::|||:::||||||||||      :::|||||||||||||||||||||:::::::::|||
Db       1220 AGTTACCCTTCTCGAAAAGAAGTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGT 1279

Qy        379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
              |||:::||||||||||||||||||   :::|||   ||||||||||||   |||||||||
Db       1280 GGGCAGACAGTGGCGCTGGTTGGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTG 1339

Qy        399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
              :::||||||||||||||||||   ||||||   :::   :::|||      |||||||||
Db       1340 ATGCAGAGGCTCTATGACCCCACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGG 1399

Qy        419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
                 :::||||||||||||   |||:::   ||||||||||||||||||||||||||||
Db       1400 ACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTG 1459

Qy        439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
```

```
Db   1460 TTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAG 1519

Qy    459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
Db   1520 ATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTTATCATGAAACTACCTAATAAA 1579

Qy    479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db   1580 TTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATC 1639

Qy    499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db   1640 GCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCA 1699

Qy    519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 538
Db   1700 GCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGC 1759

Qy    539 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558
Db   1760 CGGACTACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCCGATGTCATTGCT 1819

Qy    559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578
Db   1820 GGTTTTGATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAG 1879

Qy    579 GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 598
Db   1880 GGCATTTACTTCAAACTTGTCACAATGCAG---ACAAGAGGAAATGAAATTGAGTTAGAA 1936

Qy    599 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 618
Db   1937 AATGCCACTGGTGAATCCAAAAGTGAAAGTGATGCCTTGGAAATG---TCTCCAAAAGAT 1993

Qy    619 IleLysSerAspPheIle------------------------------------------ 624
Db   1994 TCAGGGTCCAGTTTAATAAAAAGAAGATCAACTCGCAGGAGTATACATGCACCACAAGGC 2053

Qy    625 ---AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643
Db   2054 CAAGACAGAAAGCTTGGTACAAAAGAGGACTTGAATGAGAATGTACCTCCAGTTTCCTTC 2113

Qy    644 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 663
Db   2114 TGGAGGATTCTGAAGCTGAACTCAACTGAATGGCCTTATTTTGTGGTTGGTATATTTTGT 2173

Qy    664 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 683
Db   2174 GCTATTATAAACGGAGGCCTGCAACCAGCATTTTCAATAATATTTTCAAGGATTATAGGG 2233

Qy    684 MetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 702
Db   2234 ATCTTTACCCGAGATGAGGATCCTGAAACAAAACGACAGAATAGTAACATGTTTTCTGTA 2293

Qy    703 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722
Db   2294 TTGTTTCTAGTCCTTGGAATTATTTCTTTTATTACATTTTTCCTCCAGGGCTTCACATTT 2353

Qy    723 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 742
Db   2354 GGCAAAGCTGGGGAGATCCTCACTAAGCGGCTTCGATACATGGTTTTCAGATCCATGCTG 2413

Qy    743 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIle 762
Db   2414 AGACAGGATGTCAGCTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACAACCAGG 2473

Qy    763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 782
Db   2474 CTTGCCAATGATGCGGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTCATTACC 2533

Qy    783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 802
```

```
Db   2534 CAGAATATAGCAAATCTTGGGACAGGCATTATTATATCCTTAATCTATGGTTGGCAATTA 2593

Qy    803 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 822
Db   2594 ACACTTTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAA 2653

Qy    823 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 842
Db   2654 ATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCT 2713

Qy    843 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 862
Db   2714 ACAGAAGCCATCGAAAACTTCCGAACTGTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAA 2773

Qy    863 GlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 882
Db   2774 TACATGTATGCACAGAGTTTGCAAGTACCATACAGAAACTCTTTGAGGAAAGCACACATC 2833

Qy    883 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 902
Db   2834 TTCGGGGTCTCATTTTCTATCACCCAGGCAATGATGTATTTTTCCTATGCTGGCTGTTTC 2893

Qy    903 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 922
Db   2894 CGGTTTGGTGCCTACTTGGTGGCAAATGAGTTCATGAACTTTCAGGATGTTCTTTTGGTA 2953

Qy    923 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 942
Db   2954 TTCTCAGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAGGTCAGTTCATTTGCTCCTGAC 3013

Qy    943 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 962
Db   3014 TATGCCAAAGCCAAAGTATCAGCAGCCCACGTCATCATGATCATTGAAAAAAGCCCTCTG 3073

Qy    963 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 982
Db   3074 ATTGACAGCTACAGCCCTCACGGCCTCAAGCCAAATACGTTGGAAGGAAATGTGACATTT 3133

Qy    983 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002
Db   3134 AATGAGGTCGTGTTCAACTATCCCACTCGACCAGACATCCCCGTGCTCCAGGGGCTGAGC 3193

Qy   1003 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 1022
Db   3194 CTCGAGGTGAAGAAGGGCCAGACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGC 3253

Qy   1023 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1042
Db   3254 ACAGTTGTTCAGCTCCTAGAGCGCTTCTATGACCCCTTGGCTGGTTCAGTGCTAATTGAT 3313

Qy   1043 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 1062
Db   3314 GGCAAAGAGATAAAGCACCTGAATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCT 3373

Qy   1063 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082
Db   3374 CAGGAGCCCATCCTGTTTGACTGCAGCATTGCCGAGAACATTGCCTATGGAGACAACAGC 3433

Qy   1083 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1102
Db   3434 CGGGTCGTATCACATGAAGAGATTATGCAGGCAGCCAAGGAGGCCAACATACACCACTTC 3493

Qy   1103 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122
Db   3494 ATCGAGACACTCCCTGAGAAATACAACACCAGAGTAGGAGACAAAGGAACCCAGCTCTCT 3553

Qy   1123 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 1142
Db   3554 GGTGGCCAGAAACAGCGCATTGCCATAGCTCGCGCTCTTGTTAGACAGCCTCATATTTTG 3613

Qy   1143 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 1162
Db   3614 CTTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCC 3673
```

```
Qy    1163 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 1182
           ||||||||||||||    |||||||||||:::|||:::  |||||||||||    |||
Db    3674 CTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATCGCCCACCGCTTGTCCACCATC 3733

Qy    1183 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 1202
           |||||||||||||||||||||       ||||||||:::||||||   ||||||||
Db    3734 CAGAATGCAGATTTAATAGTGGTGTTTCAGAATGGCAAAGTCAAGGAGCATGGCACACAT 3793

Qy    1203 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
           |||:::||||||      :::  |||||||||   :::|||:::   |||:::
Db    3794 CAACAGCTGCTGGCCCAGAAAGGCATCTATTTTTCCATGGTCAGTGTCCAGGCT 3847
```

```
RESULT 13
AX105080
LOCUS       AX105080                 4279 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 24 from Patent WO0123540.
ACCESSION   AX105080
VERSION     AX105080.1  GI:13921230
KEYWORDS    .
SOURCE      dog.
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 4279)
  AUTHORS   Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and
            Patten,C.J.
  TITLE     P-glycoproteins and uses thereof
  JOURNAL   Patent: WO 0123540-A 24 05-APR-2001;
            GENTEST CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..4279
                     /organism="Canis familiaris"
                     /db_xref="taxon:9615"
     CDS             17..3862
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC37722.1"
                     /db_xref="GI:13921231"
                     /translation="MDPEGGRKGSAEKNFWKMGKKSKKKEKKEKKPTVSTFAMFRYSN
                     WLDRLYMLVGTMAAIIHGAALPLMMLVFGNMTDSFANAGISRNKTFPVIINESITNNT
                     QHFINHLEEEMTTYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQILKIRKQFFHAIMRQ
                     EIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSIATFFTGFIVGFTRGWKLTL
                     VILAISPVLGLSAAIWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFGGQKKEL
                     ERYNKNLEEAKGIGIKKAITANISIGAAFLLIYASYALAFWYGTSLVLSSEYSIGQVL
                     TVFFSVLIGAFSIGQASPSIEAFANARGAAYEIFKIIDNKPSIDSYSKSGHKPDNIKG
                     NLEFKNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQRLYDPTD
                     GMVCIDGQDIRTINVRHLREITGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKE
                     ANAYDFIMKLPNKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTE
                     SEAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELMKEKGIY
                     FKLVTMQTRGNEIELENATGESKSESDALEMSPKDSGSSLIKRRSTRRSIHAPQGQDR
                     KLGTKEDLNENVPPVSFWRILKLNSTEWPYFVVGIFCAIINGGLQPAFSIIFSRIIGI
                     FTRDEDPETKRQNSNMFSVLFLVLGIISFITFFLQGFTFGKAGEILTKRLRYMVFRSM
                     LRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAVITQNIANLGTGIIISLIYG
                     WQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRTVVSLTR
                     EQKFEYMYAQSLQVPYRNSLRKAHIFGVSFSITQAMMYFSYAGCFRFGAYLVANEFMN
                     FQDVLLVFSAIVFGAMAVGQVSSFAPDYAKAKVSAAHVIMIIEKSPLIDSYSPHGLKP
                     NTLEGNVTFNEVVFNYPTRPDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERF
                     YDPLAGSVLIDGKEIKHLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSHEE
                     IMQAAKEANIHHFIETLPEKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEA
                     TSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGKVKEHGTHQQL
                     LAQKGIYFSMVSVQAGAKR"
BASE COUNT     1296 a    833 c   1008 g   1142 t
ORIGIN

Alignment Scores:
Pred. No.:              4.45e-227      Length:         4279
Score:                  3530.50        Matches:        682
Percent Similarity:     73.13%         Conservative:   238
Best Local Similarity:  54.21%         Mismatches:     269
Query Match:            57.32%         Indels:         69
DB:                     6              Gaps:           9
```

```
US-09-873-409-5 (1-1222) x AX105080 (1-4279)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||::::::|||   :::|||::::::::::::||||||   |||||||||||    |||
Db     167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
           |||   |||   |||:::|||:::       :::       :::
Db     227 GTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy      35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                                :::      ||||||    |||
Db     287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy      45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                          ||||||   |||||||||                |||:::
Db     347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy      60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
              |||||||||:::|||   |||   :::   |||   ||||||        :::||||||
Db     407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy      80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
           |||||||||||||||:::::::::     |||:::||||||||||||||        |||:::
Db     467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy     100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
           ||||||||||||||||||:::|||   ||||:::   ||||||::::::|||||||||||
Db     527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy     119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
           ||||||   :::   ||||||:::::::::||||||        |||      ||||||
Db     587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy     139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
              :::||||||||||||||||||||||    |||:::   ||||||::::::      |||
Db     647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy     159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
           ||||||   ::::::::::::   |||   |||   |||||||||   ||||||:::
Db     707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy     179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
           |||||||||||||||||||||||||||||||::::::||||||||||||||||||||||
Db     767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy     199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
              |||:::|||||||||:::
Db     827 GGACAAAAGAAAGAACTTGAA--------------------------------------- 847

Qy     219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
                                                   ||||||   :::||||||
Db     848 ------------------------------------------AGGTACAACAAAAATTTA 865

Qy     239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
           ::::::||||||      |||||||||:::   |||   :::   :::|||:::||||||
Db     866 GAAGAAGCTAAAGGAATTGGGATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCC 925

Qy     259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
              :::      :::      :::|||   ||||||||||||||||||||||||||||:::
Db     926 GCTTTCTTATTGATCTATGCATCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTC 985

Qy     279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
           |||:::   |||      |||:::||||||   ||||||   |||||||||||||:::
Db     986 CTCTCCAGTGAA------TATTCTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTA 1039

Qy     299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
              ::::::   ||||||   |||   |||      |||   ||||||   ||||||
Db    1040 ATTGGGGCTTTTAGTATTGGACAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGA 1099

Qy     319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
           |||||||||:::   ||||||::::::||||||   |||||||||||||||::::::|||
Db    1100 GGAGCAGCTTATGAAATCTTCAAGATAATTGACAATAAACCAAGCATTGACAGCTATTCG 1159
```

```
Qy     339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
Db    1160 AAGAGTGGACATAAACCAGATAATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTC 1219

Qy     359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
Db    1220 AGTTACCCTTCTCGAAAAGAAGTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGT 1279

Qy     379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
Db    1280 GGGCAGACAGTGGCGCTGGTTGGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTG 1339

Qy     399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
Db    1340 ATGCAGAGGCTCTATGACCCCACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGG 1399

Qy     419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db    1400 ACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTG 1459

Qy     439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
Db    1460 TTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAG 1519

Qy     459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
Db    1520 ATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTTATCATGAAACTACCTAATAAA 1579

Qy     479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db    1580 TTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATC 1639

Qy     499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db    1640 GCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCA 1699

Qy     519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 538
Db    1700 GCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGC 1759

Qy     539 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558
Db    1760 CGGACTACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCCGATGTCATTGCT 1819

Qy     559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578
Db    1820 GGTTTTGATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAG 1879

Qy     579 GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 598
Db    1880 GGCATTTACTTCAAACTTGTCACAATGCAG---ACAAGAGGAAATGAAATTGAGTTAGAA 1936

Qy     599 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 618
Db    1937 AATGCCACTGGTGAATCCAAAAGTGAAAGTGATGCCTTGGAAATG---TCTCCAAAAGAT 1993

Qy     619 IleLysSerAspPheIle------------------------------------------ 624
Db    1994 TCAGGGTCCAGTTTAATAAAAAGAAGATCAACTCGCAGGAGTATACATGCACCACAAGGC 2053

Qy     625 ---AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643
Db    2054 CAAGACAGAAAGCTTGGTACAAAAGAGGACTTGAATGAGAATGTACCTCCAGTTTCCTTC 2113

Qy     644 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 663
Db    2114 TGGAGGATTCTGAAGCTGAACTCAACTGAATGGCCTTATTTTGTGGTTGGTATATTTTGT 2173

Qy     664 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 683
Db    2174 GCTATTATAAACGGAGGCCTGCAACCAGCATTTTCAATAATATTTTCAAGGATTATAGGG 2233

Qy     684 MetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 702
Db    2234 ATCTTTACCCGAGATGAGGATCCTGAAACAAAACGACAGAATAGTAACATGTTTTCTGTA 2293

Qy     703 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722
Db    2294 TTGTTTCTAGTCCTTGGAATTATTTCTTTTATTACATTTTTCCTCCAGGGCTTCACATTT 2353

Qy     723 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 742
Db    2354 GGCAAAGCTGGGGAGATCCTCACTAAGCGGCTTCGATACATGGTTTTCAGATCCATGCTG 2413

Qy     743 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIle 762
Db    2414 AGACAGGATGTCAGCTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACAACCAGG 2473

Qy     763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 782
Db    2474 CTTGCCAATGATGCGGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTCATTACC 2533

Qy     783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 802
Db    2534 CAGAATATAGCAAATCTTGGGACAGGCATTATTATATCCTTAATCTATGGTTGGCAATTA 2593

Qy     803 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 822
Db    2594 ACACTTTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAA 2653

Qy     823 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 842
Db    2654 ATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCT 2713

Qy     843 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 862
Db    2714 ACAGAAGCCATCGAAAACTTCCGAACTGTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAA 2773

Qy     863 GlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 882
Db    2774 TACATGTATGCACAGAGTTTGCAAGTACCATACAGAAACTCTTTGAGGAAAGCACACATC 2833

Qy     883 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 902
Db    2834 TTCGGGGTCTCATTTTCTATCACCCAGGCAATGATGTATTTTTCCTATGCTGGCTGTTTC 2893

Qy     903 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 922
Db    2894 CGGTTTGGTGCCTACTTGGTGGCAAATGAGTTCATGAACTTTCAGGATGTTCTTTTGGTA 2953

Qy     923 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 942
Db    2954 TTCTCAGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAGGTCAGTTCATTTGCTCCTGAC 3013

Qy     943 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 962
Db    3014 TATGCCAAAGCCAAAGTATCAGCAGCCCACGTCATCATGATCATTGAAAAAAGCCCTCTG 3073

Qy     963 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 982
Db    3074 ATTGACAGCTACAGCCCTCACGGCCTCAAGCCAAATACGTTGGAAGGAAATGTGACATTT 3133

Qy     983 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002
Db    3134 AATGAGGTCGTGTTCAACTATCCCACTCGACCAGACATCCCCGTGCTCCAGGGGCTGAGC 3193

Qy    1003 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 1022
Db    3194 CTCGAGGTGAAGAAGGGCCAGACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGC 3253

Qy    1023 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1042
Db    3254 ACAGTTGTTCAGCTCCTAGAGCGCTTCTATGACCCCTTGGCTGGTTCAGTGCTAATTGAT 3313

Qy    1043 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 1062
```

```
               |||   :::   |||   ||||||||||||||||||||:::    :::   ||||||
Db      3314 GGCAAAGAGATAAAGCACCTGAATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCT 3373

Qy      1063 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082
             ||||||||:::||||||:::||||||||||||||||||||||||||||||||||||||||
Db      3374 CAGGAGCCCATCCTGTTTGACTGCAGCATTGCCGAGAACATTGCCTATGGAGACAACAGC 3433

Qy      1083 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1102
             ||||||||      :::||||||   :::||||||      |||||||||||     |||
Db      3434 CGGGTCGTATCACATGAAGAGATTATGCAGGCAGCCAAGGAGGCCAACATACACCACTTC 3493

Qy      1103 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122
             ||||||   ||||||||||||||||||||:::||||||   ||||||   |||||||||
Db      3494 ATCGAGACACTCCCTGAGAAATACAACACCAGAGTAGGAGACAAAGGAACCCAGCTCTCT 3553

Qy      1123 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 1142
             ||||||||||||||||||:::|||||||||||||||||||:::::::::|||   ||||||
Db      3554 GGTGGCCAGAAACAGCGCATTGCCATAGCTCGCGCTCTTGTTAGACAGCCTCATATTTTG 3613

Qy      1143 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 1162
             |||||||||||||||||||||||||||||||     :::|||||||||||||||||   |||
Db      3614 CTTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCC 3673

Qy      1163 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 1182
             ||||||||||||||||||||||||||||:::   ||||||||||||      |||
Db      3674 CTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATCGCCCACCGCTTGTCCACCATC 3733

Qy      1183 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 1202
             ||||||||||||||||||||||||      ||||||||||:::||||||   |||||||||
Db      3734 CAGAATGCAGATTTAATAGTGGTGTTTCAGAATGGCAAAGTCAAGGAGCATGGCACACAT 3793

Qy      1203 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
             |||:::||||||     :::   |||||||||   :::|||:::   |||:::
Db      3794 CAACAGCTGCTGGCCCAGAAAGGCATCTATTTTTCCATGGTCAGTGTCCAGGCT 3847

RESULT 14
AX391099
LOCUS       AX391099                4646 bp    DNA     linear   PAT 19-MAR-2002
DEFINITION  Sequence 1 from Patent WO0210205.
ACCESSION   AX391099
VERSION     AX391099.1  GI:19584228
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ruth,A. and Roninson,I.
  TITLE     Mutations of the mdr1 p-glycoprotein that improve its ability to
            confer resistance to chemotherapeutic drugs
  JOURNAL   Patent: WO 0210205-A 1 07-FEB-2002;
            THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)
FEATURES             Location/Qualifiers
     source          1..4646
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             425..4267
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD28599.1"
                     /db_xref="GI:19584229"
                     /translation="MDLEGDRNGGAKKKNFFKLNNKSEKDKKEKKPTVSVFSMFRYSN
                     WLDKLYMVVGTLAAIIHGAGLPLMMLVFGEMTDIFANAGNLEDLMSNITNRSDINDTG
                     FFMNLEEDMTRYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQIHKIRKQFFHAIMRQEI
                     GWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSMATFFTGFIVGFTRGWKLTLVI
                     LAISPVLGLSAAVWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFGGQKKELER
                     YNKNLEEAKRIGIKKAITANISIGAAFLLIYASYALAFWYGTTLVLSGEYSIGQVLTV
                     FFSVLIGAFSVGQASPSIEAFANARGAAYEIFKIIDNKPSIDSYSKSGHKPDNIKGNL
                     EFRNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQRLYDPTEGM
                     VSVDGQDIRTINVRFLREIIGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKEAN
                     AYDFIMKLPHKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTESE
                     AVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELMKEKGIYFK
                     LVTMQTAGNEVELENAADESKSEIDALEMSSNDSRSSLIRKRSTRRSVRGSQAQDRKL
                     STKEALDESIPPVSFWRIMKLNLTEWPYFVVGVFCAIINGGLQPAFAIIFSKIIGVFT
                     RIDDPETKRQNSNLFSLLFLALGIISFITFFLQGFTFGKAGEILTKRLRYMVFRSMLR
                     QDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAVITQNIANLGTGIIISFIYGWQ
                     LTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRTVVSLTQEQ
                     KFEHMYAQSLQVPYRNSLRKAHIFGITFSFTQAMMYFSYAGCFRFGAYLVAHKLMSFE
                     DVLLVFSAVVFGAMAVGQVSSFAPDYAKAKISAAHIIMIIEKTPLIDSYSTEGLMPNT
                     LEGNVTFGEVVFNYPTRPDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERFYD
                     PLAGKVLLDGKEIKRLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSQEEIV
                     RAAKEANIHAFIESLPNKYSTKVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEATS
                     ALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGRVKEHGTHQQLLA
                     QKGIYFSMVSVQAGTKRQ"
BASE COUNT     1372 a    892 c   1130 g   1252 t
ORIGIN

Alignment Scores:
Pred. No.:              4.95e-227     Length:          4646
Score:                  3530.50       Matches:         682
Percent Similarity:     74.00%        Conservative:    246
Best Local Similarity:  54.39%        Mismatches:      263
Query Match:            57.32%        Indels:          63
DB:                     6             Gaps:            10

US-09-873-409-5 (1-1222) x AX391099 (1-4646)

Qy         1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
             |||::::::|||   ||||||::::::::::::||||||   |||||||||||    |||
Db       575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy        21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
             |||   |||||||||:::|||                          :::|||:::|||
Db       635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy        32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe----------------- 43
                      ::::::       |||      ||||||
Db       695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy        44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                   |||      ||||||   |||||||||            |||:::   :::
Db       755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy        62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
             ||||||:::|||   |||   :::   |||   ||||||      :::||||||||||||
Db       815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy        82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
             ||||||||:::::::::   |||:::||||||||||||||||||      |||:::||||||
Db       875 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy       102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
             ||||||||||||:::|||   |||:::   ||||||::::::|||||||||||||||||||
Db       935 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGGAATTGGTGACAAAATT 994

Qy       121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
                :::   ||||||:::|||:::||||||            |||   ||||||   :::
Db       995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy       141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
             ||||||||||||||||||   |||:::   ||||||::::::         |||||||||
Db      1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy       161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
                ::::::::::::   |||   |||   |||||||||   ||||||:::||||||
Db      1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy       181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
             |||||||||||||||||||||||::::::||||||||||||||||||||||      |||
Db      1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy       201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
             :::||||||||||:::
Db      1235 AAGAAAGAACTTGAA--------------------------------------------- 1249

Qy       221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
```

```
Db   1250 -----------------------------------AGGTACAACAAAAATTTAGAAGAA 1273

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db   1274 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 1333

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db   1334 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 1393

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db   1394 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1447

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db   1448 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1507

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   1508 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1567

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   1568 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1627

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   1628 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1687

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   1688 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1747

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   1748 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1807

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   1808 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1867

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   1868 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1927

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   1928 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1987

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   1988 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 2047

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   2048 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 2107

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 540
Db   2108 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACC 2167

Qy    541 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 560
Db   2168 ACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTC 2227

Qy    561 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 580
Db   2228 GATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATT 2287

Qy    581 TyrTyrSerLeuValMetSerGln------------------AspIleLysLysAlaAsp 594
Db   2288 TACTTCAAACTTGTCACAATGCAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCT 2347

Qy    595 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHis 614
Db   2348 GATGAATCCAAAAGTGAAATTGATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGT 2407

Qy    615 SerVal------------------LysSerIleLys---SerAspPheIleAspLysAla 627
Db   2408 CTAATAAGAAAAAGATCAACTCGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAG 2467

Qy    628 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 647
Db   2468 CTTAGTACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATG 2527

Qy    648 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 667
Db   2528 AAGCTAAATTTAACTGAATGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAAT 2587

Qy    668 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 687
Db   2588 GGAGGCCTGCAACCAGCATTTGCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGA 2647

Qy    688 ---AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIle 706
Db   2648 ATTGATGATCCTGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCC 2707

Qy    707 LeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGly 726
Db   2708 CTTGGAATTATTTCTTTTATTACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGA 2767

Qy    727 GluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIle 746
Db   2768 GAGATCCTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTG 2827

Qy    747 AlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAsp 766
Db   2828 AGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGAT 2887

Qy    767 IleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThr 786
Db   2888 GCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCA 2947

Qy    787 AsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIle 806
Db   2948 AATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTC 3007

Qy    807 LeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGly 826
Db   3008 TTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGA 3067

Qy    827 PheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeu 846
Db   3068 CAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATA 3127

Qy    847 GluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGlu 866
Db   3128 GAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCT 3187

Qy    867 GluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCys 886
Db   3188 CAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACA 3247

Qy    887 TyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAla 906
Db   3248 TTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCC 3307

Qy    907 TyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIle 926
Db   3308 TACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTT 3367

Qy    927 AlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAla 946
Db   3368 GTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCC 3427
```

```
Qy     947 LysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArg 966
Db    3428 AAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTAC 3487

Qy     967 SerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSer 986
Db    3488 AGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTA 3547

Qy     987 PhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGlu 1006
Db    3548 TTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAG 3607

Qy    1007 ArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGln 1026
Db    3608 AAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAG 3667

Qy    1027 LeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAla 1046
Db    3668 CTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATA 3727

Qy    1047 LysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProVal 1066
Db    3728 AAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATC 3787

Qy    1067 LeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValPro 1086
Db    3788 CTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCA 3847

Qy    1087 LeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeu 1106
Db    3848 CAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTG 3907

Qy    1107 ProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLys 1126
Db    3908 CCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAA 3967

Qy    1127 GlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGlu 1146
Db    3968 CAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAA 4027

Qy    1147 AlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAla 1166
Db    4028 GCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCC 4087

Qy    1167 ArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAsp 1186
Db    4088 AGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGAC 4147

Qy    1187 LeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeu 1206
Db    4148 TTAATAGTGGTGTTTCAGAATGGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTG 4207

Qy    1207 ArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db    4208 GCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 4249
```

```
RESULT 15
AF269224
LOCUS       AF269224                4045 bp    DNA     linear   SYN 04-JUL-2000
DEFINITION  Synthetic construct Canis familiaris his-tagged-multidrug
            resistance glycoprotein gene, complete cds.
ACCESSION   AF269224
VERSION     AF269224.1  GI:8926216
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4045)
  AUTHORS   Roulet,A.C., Puel,O., Gesta,S., Alvinerie,M. and Pineau,T.J.
  TITLE     Analysis of dog MDR1 p-glycoprotein
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4045)
  AUTHORS   Roulet,A.C., Puel,O., Gesta,S., Alvinerie,M. and Pineau,T.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-MAY-2000) Pharmacology, INRA, 180 Chemin de
            Tournefeuille, BP3, Toulouse 31931, France
FEATURES             Location/Qualifiers
     source          1..4045
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="original mRNA isolated from Beagle dog colon"
     CDS             101..3964
                     /codon_start=1
                     /transl_table=11
                     /product="his-tagged-multidrug resistance glycoprotein
                     MDR1"
                     /protein_id="AAF81747.1"
                     /db_xref="GI:8926217"
                     /translation="MHHHHHHDPEGGRKGSAEKNFWKMGKKSKKNEKKEKKPTVSTFA
                     MFRYSNWLDRLYMLVGTMAAIIHGAALPLMMLVFGNMTDSFANAGISRNKTFPVIINE
                     SITNNTQHFINHLEEEMTTYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQILKIRKQFF
                     HAIMRQEIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSIATFFTGFIVGFTR
                     GWKLTLVILAISPVLGLSAAIWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFG
                     GQKKELERYNKNLEEAKRIGIKKAITANISIGAAFLLIYASYALAFWYGTSLVLSSEY
                     SIGQVLTVFFSVLIGAFSIGQASPSIEAFANARGAAYEIFKIIDNKPSIDSYSKSGHK
                     PDNIKGNLEFKNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQR
                     LYDPTDGMVCIDGQDIRTINVRHLREITGVVSQEPVLFATTIAENIRYGRENVTMDEI
                     EKAVKEANAYDFIMKLPNKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEAT
                     SALDTESEAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELM
                     KEKGIYFKLVTMQTRGNEIELENATGESKSESDALEMSPKDSGSSLIKRRSTRRSIHA
                     PQGQDRKLGTKEDLNENVPPVSFWRILKLNSTEWPYFVVGIFCAIINGGLQPAFSIIF
                     SRIIGIFTRDEDPETKRQNSNMFSVLFLVLGIISFITFFLQGFTFGKAGEILTKRLRY
                     MVFRSMLRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAVITQNIANLGTGII
                     ISLIYGWQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRT
                     VVSLTREQKFEYMYAQSLQVPYRNSLRKAHIFGVSFSITQAMMYFSYAGCFRFGAYLV
                     ANEFMNFQDVLLVFSAIVFGAMAVGQVSSFAPDYAKAKVSAAHVIMIIEKSPLIDSYS
                     PHGLKPNTLEGNVTFNEVVFNYPTRPDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVV
                     QLLERFYDPLAGSVLIDGKEIKHLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSR
                     VVSHEEIMQAAKEANIHHFIETLPEKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHI
                     LLLDEATSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGKVKEH
                     GTHQQLLAQKGIYFSMISVQAGAKR"
BASE COUNT     1171 a    850 c    986 g   1038 t
ORIGIN
```

```
Alignment Scores:
Pred. No.:              5.64e-227     Length:          4045
Score:                  3528.50       Matches:         681
Percent Similarity:     73.13%        Conservative:    239
Best Local Similarity:  54.13%        Mismatches:      269
Query Match:            57.29%        Indels:          69
DB:                     12            Gaps:            9
```

US-09-873-409-5 (1-1222) x AF269224 (1-4045)

```
Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db     269 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 328

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
Db     329 GTTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 388

Qy      35 --------------------ValGlnThrAsnThrTyrSerPhePheArg---------- 44
Db     389 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 448

Qy      45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db     449 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 508

Qy      60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db     509 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 568

Qy      80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
```

```
Db    569 AAACAATTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 628

Qy    100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db    629 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 688

Qy    119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db    689 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 748

Qy    139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db    749 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 808

Qy    159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db    809 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 868

Qy    179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db    869 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 928

Qy    199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
Db    929 GGACAAAAGAAAGAACTTGAA--------------------------------------- 949

Qy    219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
Db    950 ------------------------------------------AGGTACAACAAAAATTTA 967

Qy    239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
Db    968 GAAGAAGCTAAAAGAATTGGGATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCC 1027

Qy    259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
Db   1028 GCTTTCTTATTGATCTATGCATCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTC 1087

Qy    279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
Db   1088 CTCTCCAGTGAA------TATTCTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTA 1141

Qy    299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
Db   1142 ATTGGGGCTTTTAGTATTGGACAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGA 1201

Qy    319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
Db   1202 GGAGCAGCTTATGAAATCTTCAAGATAATTGACAATAAACCAAGCATTGACAGCTATTCG 1261

Qy    339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
Db   1262 AAGAGTGGACATAAACCAGATAATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTC 1321

Qy    359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
Db   1322 AGTTACCCTTCTCGAAAAGAAGTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGT 1381

Qy    379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
Db   1382 GGGCAGACAGTGGCGCTGGTTGGAAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTG 1441

Qy    399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
Db   1442 ATGCAGAGGCTCTATGACCCCACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGG 1501

Qy    419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db   1502 ACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTG 1561

Qy    439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
Db   1562 TTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAG 1621

Qy    459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
Db   1622 ATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTTATCATGAAACTACCTAATAAA 1681

Qy    479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db   1682 TTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATC 1741

Qy    499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db   1742 GCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCA 1801

Qy    519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 538
Db   1802 GCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGC 1861

Qy    539 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 558
Db   1862 CGGACTACCATTGTGATAGCTCATCGTTTGTCTACAGTTCGTAATGCCGATGTCATTGCT 1921

Qy    559 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 578
Db   1922 GGTTTTGATGATGGAGTCATTGTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAG 1981

Qy    579 GlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 598
Db   1982 GGCATTTACTTCAAACTTGTCACAATGCAG---ACAAGAGGAAATGAAATTGAGTTAGAA 2038

Qy    599 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 618
Db   2039 AATGCCACTGGTGAATCCAAAAGTGAAAGTGATGCCTTGGAAATG---TCTCCAAAAGAT 2095

Qy    619 IleLysSerAspPheIle------------------------------------------ 624
Db   2096 TCAGGGTCCAGTTTAATAAAAAGAAGATCAACTCGCAGGAGTATACATGCACCACAAGGC 2155

Qy    625 ---AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeu 643
Db   2156 CAAGACAGAAAGCTTGGTACAAAAGAGGACTTGAATGAGAATGTACCTCCAGTTTCCTTC 2215

Qy    644 LeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAla 663
Db   2216 TGGAGGATTCTGAAGCTGAACTCAACTGAATGGCCTTATTTTGTGGTTGGTATATTTTGT 2275

Qy    664 SerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThr 683
Db   2276 GCTATTATAAACGGAGGCCTGCAACCAGCATTTTCAATAATATTTTCAAGGATTATAGGG 2335

Qy    684 MetPhe---GlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMet 702
Db   2336 ATCTTTACCCGAGATGAGGATCCTGAAACAAAACGACAGAATAGTAACATGTTTTCTGTA 2395

Qy    703 IlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyr 722
Db   2396 TTGTTTCTAGTCCTTGGAATTATTTCTTTTATTACATTTTTCCTCCAGGGCTTCACATTT 2455

Qy    723 GlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeu 742
Db   2456 GGCAAAGCTGGGGAGATCCTCACTAAGCGGCTTCGATACATGGTTTTCAGATCCATGCTG 2515

Qy    743 TyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIle 762
Db   2516 AGACAGGATGTCAGCTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACAACCAGG 2575

Qy    763 LeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThr 782
Db   2576 CTTGCCAATGATGCGGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTCATTACC 2635

Qy    783 GlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMet 802
Db   2636 CAGAATATAGCAAATCTTGGGACAGGCATTATTATATCCTTAATCTATGGTTGGCAATTA 2695
```

```
Qy     803 ThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAla 822
Db    2696 ACACTTTTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAA 2755

Qy     823 AlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAla 842
Db    2756 ATGTTGTCTGGACAAGCACTGAAAGATAAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCT 2815

Qy     843 ThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGlu 862
Db    2816 ACAGAAGCCATCGAAAACTTCCGAACTGTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAA 2875

Qy     863 GlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIle 882
Db    2876 TACATGTATGCACAGAGTTTGCAAGTACCATACAGAAACTCTTTGAGGAAAGCACACATC 2935

Qy     883 IleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPhe 902
Db    2936 TTCGGGGTCTCATTTTCTATCACCCAGGCAATGATGTATTTTTCCTATGCTGGCTGTTTC 2995

Qy     903 ArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleVal 922
Db    2996 CGGTTTGGTGCCTACTTGGTGGCAAATGAGTTCATGAACTTTCAGGATGTTCTTTTGGTA 3055

Qy     923 PheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGlu 942
Db    3056 TTCTCAGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAGGTCAGTTCATTTGCTCCTGAC 3115

Qy     943 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 962
Db    3116 TATGCCAAAGCCAAAGTATCAGCAGCCCACGTCATCATGATCATTGAAAAAAGCCCTCTG 3175

Qy     963 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 982
Db    3176 ATTGACAGCTACAGCCCTCACGGCCTCAAGCCAAATACGTTGGAAGGAAATGTGACATTT 3235

Qy     983 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 1002
Db    3236 AATGAGGTCGTGTTCAACTATCCCACTCGACCAGACATCCCCGTGCTCCAGGGGCTGAGC 3295

Qy    1003 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 1022
Db    3296 CTCGAGGTGAAGAAGGGCCAGACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGC 3355

Qy    1023 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1042
Db    3356 ACAGTTGTTCAGCTCCTAGAGCGCTTCTATGACCCCTTGGCTGGTTCAGTGCTAATTGAT 3415

Qy    1043 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 1062
Db    3416 GGCAAAGAGATAAAGCACCTGAATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCT 3475

Qy    1063 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1082
Db    3476 CAGGAGCCCATCCTGTTTGACTGCAGCATTGCCGAGAACATTGCCTATGGAGACAACAGC 3535

Qy    1083 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1102
Db    3536 CGGGTCGTATCACATGAAGAGATTATGCAGGCAGCCAAGGAGGCCAACATACACCACTTC 3595

Qy    1103 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1122
Db    3596 ATCGAGACACTCCCTGAGAAATACAACACCAGAGTAGGAGACAAAGGAACCCAGCTCTCT 3655

Qy    1123 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeu 1142
Db    3656 GGTGGCCAGAAACAGCGCATTGCCATAGCTCGCGCTCTTGTTAGACAGCCTCATATTTTG 3715

Qy    1143 LeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAla 1162
Db    3716 CTTTTGGATGAAGCTACATCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCC 3775

Qy    1163 LeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIle 1182
Db    3776 CTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATCGCCCACCGCTTGTCCACCATC 3835

Qy    1183 GlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHis 1202
Db    3836 CAGAATGCAGATTTAATAGTGGTGTTTCAGAATGGCAAAGTCAAGGAGCATGGCACACAT 3895

Qy    1203 GlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1220
Db    3896 CAACAGCTGCTGGCTCAGAAAGGCATCTATTTTTCCATGATCAGTGTCCAGGCT 3949
```

Search completed: March 31, 2003, 02:18:16
Job time : 16788.1 secs

```
                 GenCore version 5.1.4_p5_4578
        Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:        March 31, 2003, 13:52:37 ; Search time 434.971 Seconds
                                          (without alignments)
                                          2337.888 Million cell updates/sec

Title:          US-09-873-409-6
Perfect score:  6012
Sequence:       1 MILGILASLVNGACLPLMPL.........QELLRNRDIYFKLVNAQSVQ 1195

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters:      1148742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115424_19290/app_query.fasta_1.7544
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09873409_@CGN_1_1_960_@runat_27032003_115424_19290
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Published_Applications_NA:*
                1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
                2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
                3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
                4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
                5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
                6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
                7:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
                8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
                9:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
                10:  /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
                11:  /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
                12:  /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
                13:  /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
                14:  /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		6012	100.0	3621	10	US-09-873-409-14	Sequence 14, Appl
2		5988.5	99.6	3702	10	US-09-873-409-13	Sequence 13, Appl
3		5159.5	85.8	3177	10	US-09-873-409-12	Sequence 12, Appl
4		4561.5	75.9	2856	10	US-09-873-409-10	Sequence 10, Appl
5		3554	59.1	3860	10	US-09-866-866A-1	Sequence 1, Appli
6		3549	59.0	4317	9	US-10-044-671-1	Sequence 1, Appli
7		3545	59.0	3860	10	US-09-866-866A-3	Sequence 3, Appli
8		3545	59.0	8630	10	US-09-306-417-1	Sequence 1, Appli
9		3545	59.0	8630	10	US-09-306-417-2	Sequence 2, Appli
10		3540.5	58.9	4425	10	US-09-769-097-3	Sequence 3, Appli
11		3538.5	58.9	4369	10	US-09-769-097-1	Sequence 1, Appli
12		3537.5	58.8	4643	9	US-10-072-621-2	Sequence 2, Appli
13		3526.5	58.7	4189	10	US-09-866-866A-5	Sequence 5, Appli
14		3522.5	58.6	4788	10	US-09-866-866A-7	Sequence 7, Appli
15		3504.5	58.3	3924	10	US-09-880-107-2299	Sequence 2299, Ap
16		3476.5	57.8	4254	10	US-09-917-800A-1424	Sequence 1424, Ap
17		3456	57.5	2066	10	US-09-873-409-9	Sequence 9, Appli
18		3428.5	57.0	3912	10	US-09-917-800A-1560	Sequence 1560, Ap
19		2835	47.2	5010	10	US-09-917-800A-483	Sequence 483, App
20		2552	42.4	1940	10	US-09-873-409-16	Sequence 16, Appl
21		2528.5	42.1	2021	10	US-09-873-409-15	Sequence 15, Appl
22		2380.5	39.6	4175	10	US-09-749-340-3	Sequence 3, Appli
23		2351.5	39.1	3861	9	US-09-938-842A-2237	Sequence 2237, Ap
24		2282.5	38.0	4002	9	US-09-758-828-1	Sequence 1, Appli
25		2282.5	38.0	4002	9	US-09-758-828-3	Sequence 3, Appli
26		2153	35.8	3861	9	US-09-938-842A-263	Sequence 263, App
27		2046	34.0	3512	10	US-09-749-340-7	Sequence 7, Appli
28		2010.5	33.4	4653	9	US-10-101-388-2	Sequence 2, Appli
29		1659.5	27.6	3792	9	US-09-882-694-10	Sequence 10, Appl
30		1617	26.9	2698	10	US-09-749-340-5	Sequence 5, Appli
31		1550	25.8	8036	9	US-10-101-388-1	Sequence 1, Appli
32		1538.5	25.6	2681	10	US-09-749-340-8	Sequence 8, Appli
33		1525	25.4	3999	9	US-09-882-694-9	Sequence 9, Appli
34		1102.5	18.3	1810	10	US-09-749-340-4	Sequence 4, Appli
35		1101.5	18.3	1175	10	US-09-873-409-11	Sequence 11, Appl
36		1084	18.0	7296	10	US-09-070-927A-59	Sequence 59, Appl
37		1031.5	17.2	3741	9	US-09-738-626-1251	Sequence 1251, Ap
38	c	1021	17.0	9047	10	US-09-070-927A-12	Sequence 12, Appl
39		1005	16.7	12438	10	US-09-070-927A-173	Sequence 173, App
40		941	15.7	640681	10	US-09-790-988-1	Sequence 1, Appli
41	c	902.5	15.0	3601	10	US-09-070-927A-255	Sequence 255, App
42	c	901.5	15.0	6415	10	US-09-070-927A-275	Sequence 275, App
43		781	13.0	2298	9	US-10-156-239-18	Sequence 18, Appl
44		781	13.0	2298	10	US-09-795-693-18	Sequence 18, Appl
45		781	13.0	3408	9	US-10-156-239-16	Sequence 16, Appl

ALIGNMENTS

```
RESULT 1
US-09-873-409-14
; Sequence 14, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
;   LENGTH: 3621
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-409-14

Alignment Scores:
Pred. No.:               0              Length:         3621
Score:                   6012.00        Matches:        1195
Percent Similarity:      100.00%        Conservative:   0
Best Local Similarity:   100.00%        Mismatches:     0
Query Match:             100.00%        Indels:         0
DB:                      10             Gaps:           0
```

```
US-09-873-409-6 (1-1195) x US-09-873-409-14 (1-3621)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     34 ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 93

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153

Qy     41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    154 TCTTTCTTCAGGTTGACCCTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGT 213

Qy     61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    214 TACATACAGATTTCCTTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAA 273

Qy     81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    274 CAGTTTTTTCATTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333

Qy    101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    334 GAACTTAACACTCGCATGACAGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATT 393

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAG 453

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    454 GGCTGGAAACTCACCCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA 513

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAG 633

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    634 GAGAAAGAACTTCAAAGGTATACACAGAATCTCAAAGATGCAAAGGATTTTGGCATAAAA 693

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    694 AGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAATGGAACCTAT 753

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    754 GGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGATATACCATC 813

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    814 GGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATTGGAGCAGCA 873

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    874 GTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTT 933

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    934 ATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATA 993

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    994 GAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAG 1053

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1054 ATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTC 1113

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1114 AATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGAT 1173

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1174 GGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGAC 1233

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1234 CATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATC 1293

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1294 AAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAAT 1353

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1354 GCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGA 1413

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1414 GCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAAC 1473

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1474 CCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCT 1533

Qy    501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1534 GTTCAAGCTGCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTAGCACACCGA 1593

Qy    521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1594 CTTTCTACTATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATGCTGGCGGAG 1653

Qy    541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1654 AAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTTGTGATGTCA 1713

Qy    561 GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1714 CAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAG 1773

Qy    581 ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAla 600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1774 ACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCT 1833

Qy    601 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1834 GAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTA 1893

Qy    621 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1894 AAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCTGTTCTAAAT 1953

Qy    641 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1954 GGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATGTTTGGAAAT 2013

Qy    661 AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeu 680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2014 AATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTCGTCATTTTG 2073

Qy    681 GlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGlu 700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2074 GGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGAGCAGGGGAA 2133

Qy    701 IleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAla 720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2134 ATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAGGATATTGCC 2193
```

```
Qy    721 TrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIle 740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2194 TGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCCATAGATATA 2253

Qy    741 AlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsn 760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2254 GCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAATGCAACTAAC 2313

Qy    761 MetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeu 780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2314 ATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTCCTGATTCTG 2373

Qy    781 SerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPhe 800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2374 AGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATGACTGGATTT 2433

Qy    801 AlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGlu 820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2434 GCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAG 2493

Qy    821 AsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGlu 840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2494 AATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATGTATGAAGAG 2553

Qy    841 MetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyr 860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2554 ATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTAT 2613

Qy    861 AlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyr 880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2614 GCATTCAGCCATGCCTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTTGGAGCCTAT 2673

Qy    881 LeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAla 900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2674 TTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACTGCAATTGCA 2733

Qy    901 TyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLys 920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2734 TATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCCAAAGCCAAA 2793

Qy    921 SerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSer 940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2794 TCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGACAGCCGCAGT 2853

Qy    941 GlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhe 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2854 CAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTC 2913

Qy    961 PheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArg 980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2914 TTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGA 2973

Qy    981 GlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeu 1000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2974 GGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCTGTTCAACTT 3033

Qy   1001 LeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLys 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3034 CTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTGGATGCAAAA 3093

Qy   1021 GluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeu 1040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3094 GAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAGCCTGTGCTC 3153

Qy   1041 PheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeu 1060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3154 TTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTGGTGCCATTA 3213

Qy   1061 AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuPro 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3214 GATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAAGGTCTCCCT 3273

Qy   1081 GluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGln 1100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3274 GAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAACAA 3333

Qy   1101 ArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAla 1120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3334 AGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTGGATGAGGCC 3393

Qy   1121 ThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArg 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3394 ACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGATAAAGCCAGG 3453

Qy   1141 ThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeu 1160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3454 ACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTG 3513

Qy   1161 IleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArg 1180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3514 ATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAGCTCCTGAGA 3573

Qy   1181 AsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1195
          |||||||||||||||||||||||||||||||||||||||||||||
Db   3574 AATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3618
```

```
RESULT 2
US-09-873-409-13
; Sequence 13, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
;   LENGTH: 3702
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: No. US20020037522A1e
;   LOCATION: (723)..(723)
;   OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-13

Alignment Scores:
Pred. No.:              0              Length:        3702
Score:                  5988.50        Matches:       1195
Percent Similarity:     97.79%         Conservative:  0
Best Local Similarity:  97.79%         Mismatches:    0
Query Match:            99.61%         Indels:        27
DB:                     10             Gaps:          1

US-09-873-409-6 (1-1195) x US-09-873-409-13 (1-3702)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     34 ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 93

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153

Qy     41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    154 TCTTTCTTCAGGTTGACCCTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGT 213

Qy     61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    214 TACATACAGATTTCCTTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAA 273
```

```
Qy     81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    274 CAGTTTTTCATTCAGTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333

Qy    101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    334 GAACTTAACACTCGCATGACAGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATT 393

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAG 453

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    454 GGCTGGAAACTCACCCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA 513

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAG 633

Qy    201 GluLysGluLeuGln--------------------------------------------- 205
          |||||||||||||||
Db    634 GAGAAAGAACTTCAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTTGGTTTTATTTT 693

Qy    206 ------------------------------------ArgTyrThrGlnAsnLeuLysAsp 213
                                              ||||||||||||||||||||||||
Db    694 CCCCAGTGGCTACTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGAATCTCAAAGAT 753

Qy    214 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 233
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    754 GCAAAGGATTTTGGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTAC 813

Qy    234 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 253
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    814 TTCTTTATGAATGGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAAT 873

Qy    254 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 273
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    874 GGAGAACCTGGATATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGC 933

Qy    274 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 293
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    934 AGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCT 993

Qy    294 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    994 GCCTTTCATATTTTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCT 1053

Qy    314 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 333
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1054 GGATATAAACCTGAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTAT 1113

Qy    334 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1114 CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAG 1173

Qy    354 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 373
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1174 ACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG 1233

Qy    374 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 393
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1234 AGGTTATATGATCCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1293

Qy    394 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 413
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1294 AATGTGCGGCATTATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGG 1353

Qy    414 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 433
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1354 ACCACCATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1413

Qy    434 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 453
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1414 AGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAAT 1473

Qy    454 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 473
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1474 ACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATT 1533

Qy    474 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 493
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1534 GCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTG 1593

Qy    494 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 513
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1594 GATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTGGAGAAGGCGAGCAAAGGTCGGACT 1653

Qy    514 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 533
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1654 ACAATCGTGGTAGCACACCGACTTTCTACTATTCGAAGTGCAGATTTGATTGTGACCCTA 1713

Qy    534 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 553
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1714 AAGGATGGAATGCTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTA 1773

Qy    554 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet 573
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1774 TATTATTCACTTGTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATG 1833

Qy    574 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 593
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1834 ACATATTCTACTGAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAG 1893

Qy    594 SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGlu 613
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1894 TCAGACTTCATTGACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAA 1953

Qy    614 ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGly 633
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1954 GTCTCTCTATTAAAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGG 2013

Qy    634 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 653
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2014 ACATTGGCTTCTGTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAA 2073

Qy    654 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 673
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2074 ATTATAACCATGTTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTAT 2133

Qy    674 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 693
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2134 TCCATGATATTCGTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTA 2193

Qy    694 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 713
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2194 TTTTACGGCAGAGCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCC 2253

Qy    714 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 733
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2254 ATGTTATATCAGGATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACA 2313

Qy    734 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 753
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2314 ACAATATTAGCCATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTC 2373

Qy    754 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrp 773
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2374 TTAACACAAAATGCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGG 2433

Qy    774 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 793
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   2434 GAGATGACATTCCTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAA 2493

Qy    794 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 813
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2494 ACCGCAGCAATGACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAG 2553

Qy    814 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 833
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2554 ATAGCAACTGAAGCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCC 2613

Qy    834 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 853
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2614 TTCGAGCAAATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCA 2673

Qy    854 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 873
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2674 CAGATTATTGGAAGCTGTTATGCATTCAGCCATGCCTTTATATATTTGCCTATGCAGCA 2733

Qy    874 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 893
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2734 GGGTTTCGATTTGGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTC 2793

Qy    894 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 913
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2794 ATAGTTTTTACTGCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCT 2853

Qy    914 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 933
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2854 CCTGAATATTCCAAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAA 2913

Qy    934 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 953
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2914 CCAAATATAGACAGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTA 2973

Qy    954 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 973
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2974 GAGTTTCGAGAAGTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGC 3033

Qy    974 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 993
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3034 TTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGG 3093

Qy    994 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 1013
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3094 AAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTG 3153

Qy   1014 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 1033
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3154 TTTGATGGTGTGGATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATC 3213

Qy   1034 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 1053
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3214 GTTCCTCAAGAGCCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGAC 3273

Qy   1054 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 1073
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3274 AACAGCCGTGTGGTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCAT 3333

Qy   1074 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 1093
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3334 TCTTTTATTGAAGGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAG 3393

Qy   1094 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 1113
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3394 CTTTCTGGCGGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAA 3453

Qy   1114 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 1133
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3454 ATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAG 3513

Qy   1134 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 1153
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3514 CATGCCCTTGATAAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCT 3573

Qy   1154 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 1173
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3574 GCAATTCAGAACGCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGA 3633

Qy   1174 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1193
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3634 ACTCATCAAGAGCTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCA 3693

Qy   1194 ValGln 1195
          ||||||
Db   3694 GTGCAG 3699
```

RESULT 3
US-09-873-409-12
; Sequence 12, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 3177
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1e
; LOCATION: (198)..(198)
; OTHER INFORMATION: n at position 198 represents any nucleotide (A, T, C or G)
US-09-873-409-12

Alignment Scores:

Pred. No.:	0	Length:	3177
Score:	5159.50	Matches:	1031
Percent Similarity:	97.45%	Conservative:	0
Best Local Similarity:	97.45%	Mismatches:	0
Query Match:	85.82%	Indels:	27
DB:	10	Gaps:	1

US-09-873-409-6 (1-1195) x US-09-873-409-12 (1-3177)

```
Qy    165 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 184
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCA 60

Qy    185 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 204
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAGGAGAAAGAACTT 120

Qy    205 Gln--------------------------------------------------------- 205
          |||
Db    121 CAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTTGGTTTTATTTTCCCCAGTGGCTA 180

Qy    206 ------------------------ArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 217
                                  ||||||||||||||||||||||||||||||||||||
Db    181 CTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGAATCTCAAAGATGCAAAGGATTTT 240

Qy    218 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 237
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAAT 300

Qy    238 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 257
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGA 360

Qy    258 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 277
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     361 TATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATT 420

Qy     278 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 297
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATT 480

Qy     298 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 317
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCT 540

Qy     318 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 337
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCA 600

Qy     338 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 357
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 TCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTG 660

Qy     358 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 377
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 720

Qy     378 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 397
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 CCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCAT 780

Qy     398 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 417
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 TATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGT 840

Qy     418 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 437
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 AACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGG 900

Qy     438 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 457
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGG 960

Qy     458 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 477
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTA 1020

Qy     478 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 497
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGC 1080

Qy     498 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 517
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1081 AAGTCAGCTGTTCAAGCTGCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTA 1140

Qy     518 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 537
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1141 GCACACCGACTTTCTACTATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATG 1200

Qy     538 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 557
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1201 CTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTT 1260

Qy     558 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 577
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1261 GTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACT 1320

Qy     578 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 597
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1321 GAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATT 1380

Qy     598 AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 617
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1381 GACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTA 1440

Qy     618 LysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 637
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1441 AAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCT 1500

Qy     638 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 657
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1501 GTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATG 1560

Qy     658 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 677
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1561 TTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTC 1620

Qy     678 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 697
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1621 GTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGA 1680

Qy     698 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 717
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1681 GCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAG 1740

Qy     718 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 737
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1741 GATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCC 1800

Qy     738 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 757
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1801 ATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAAT 1860

Qy     758 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 777
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1861 GCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTC 1920

Qy     778 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 797
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1921 CTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATG 1980

Qy     798 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 817
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1981 ACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAA 2040

Qy     818 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 837
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2041 GCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATG 2100

Qy     838 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 857
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2101 TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGA 2160

Qy     858 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 877
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2161 AGCTGTTATGCATTCAGCCATGCCTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTT 2220

Qy     878 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 897
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2221 GGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACT 2280

Qy     898 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 917
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2281 GCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCC 2340

Qy     918 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 937
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2341 AAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGAC 2400

Qy     938 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 957
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2401 AGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAA 2460

Qy     958 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 977
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2461 GTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGT 2520

Qy     978 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 997
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2521 ATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCT 2580
```

```
Qy     998 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 1017
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2581 GTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTG 2640

Qy    1018 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 1037
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2641 GATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAG 2700

Qy    1038 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 1057
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2701 CCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTG 2760

Qy    1058 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 1077
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2761 GTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAA 2820

Qy    1078 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 1097
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2821 GGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGC 2880

Qy    1098 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 1117
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2881 CAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTG 2940

Qy    1118 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1137
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2941 GATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGAT 3000

Qy    1138 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1157
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3001 AAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAC 3060

Qy    1158 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1177
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3061 GCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAG 3120

Qy    1178 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1195
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3121 CTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3174
```

```
RESULT 4
US-09-873-409-10
; Sequence 10, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 10
;   LENGTH: 2856
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-409-10

Alignment Scores:
Pred. No.:              0           Length:          2856
Score:                  4561.50     Matches:         912
Percent Similarity:     97.96%      Conservative:    0
Best Local Similarity:  97.96%      Mismatches:      1
Query Match:            75.87%      Indels:          18
DB:                     10          Gaps:            1

US-09-873-409-6 (1-1195) x US-09-873-409-10 (1-2856)
```

```
Qy     266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHis-PheGl 285
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     114 GTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATTGGAGCAGCAGTCCCTCATTATT-- 171

Qy     285 uThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPr 305
                                                             ||||||||||
Db     172 ----------------------------------------------GATAAGAAACC 182

Qy     305 oSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGl 325
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 CAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGA 242

Qy     325 uPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLe 345
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 ATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCT 302

Qy     345 uAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLy 365
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 GAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAA 362

Qy     365 sSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVa 385
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     363 GAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGT 422

Qy     385 lAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVa 405
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     423 GGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTGGT 482

Qy     405 lSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAs 425
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     483 TAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGAGA 542

Qy     425 pAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIl 445
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     543 TGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTTAT 602

Qy     445 eMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGl 465
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     603 CATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGG 662

Qy     465 yGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIl 485
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     663 AGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGAT 722

Qy     485 eLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLe 505
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     723 TTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACT 782

Qy     505 uGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleAr 525
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     783 GGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTAGCACACCGACTTTCTACTATTCG 842

Qy     525 gSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAl 545
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     843 AAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATGCTGGCGGAGAAAGGAGCACATGC 902

Qy     545 aGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLy 565
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     903 TGAACTAATGGCAAAACGAGGTCTATATTATTCACTTGTGATGTCACAGGATATTAAAAA 962

Qy     565 sAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPr 585
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     963 AGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCACTTCC 1022

Qy     585 oLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGl 605
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1023 TCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCCACCCA 1082

Qy     605 nSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPr 625
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1083 ATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAACAAGCC 1142

Qy     625 oGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPr 645
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1143 TGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCTGTTCTAAATGGAACTGTTCATCC 1202
```

```
Qy    645 oValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrTh 665
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1203 AGTATTTTCCATCATCTTTGCAAAAATTATAACCATGTTTGGAAATAATGATAAAACCAC 1262

Qy    665 rLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPh 685
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1263 ATTAAAGCATGATGCAGAAATTTATTCCATGATATTCGTCATTTTGGGTGTTATTTGCTT 1322

Qy    685 eValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetAr 705
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1323 TGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGAGCAGGGGAAATTTTAACGATGAG 1382

Qy    705 gLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLy 725
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1383 ATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAGGATATTGCCTGGTTTGATGAAAA 1442

Qy    725 sGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGl 745
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1443 GGAAAACAGCACAGGAGGCTTGACAACAATATTAGCCATAGATATAGCACAAATTCAAGG 1502

Qy    745 yAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVa 765
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1503 AGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAATGCAACTAACATGGGACTTTCAGT 1562

Qy    765 lIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVa 785
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1563 TATCATTTCCTTTATATATGGATGGGAGATGACATTCCTGATTCTGAGTATTGCTCCAGT 1622

Qy    785 lLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLy 805
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1623 ACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATGACTGGATTTGCCAACAAAGATAA 1682

Qy    805 sGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIl 825
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1683 GCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAGAATATACGTACTAT 1742

Qy    825 eValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGl 845
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1743 AGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATGTATGAAGAGATGCTTCAGACTCA 1802

Qy    845 nHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAl 865
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1803 ACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTATGCATTCAGCCATGC 1862

Qy    865 aPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGl 885
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1863 CTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTTGGAGCCTATTTAATTCAAGCTGG 1922

Qy    885 yArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAl 905
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1923 ACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACTGCAATTGCATATGGAGCTATGGC 1982

Qy    905 aIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHi 925
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1983 CATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCCAAAGCCAAATCGGGGGCTGCGCA 2042

Qy    925 sLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLy 945
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2043 TCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGACAGCCGCAGTCAAGAAGGGAAAAA 2102

Qy    945 sProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysAr 965
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2103 GCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTCTTCTATCCATGTCG 2162

Qy    965 gProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAl 985
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2163 CCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGC 2222

Qy    985 aPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTy 1005
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2223 ATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTA 2282

Qy   1005 rAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGl 1025
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2283 TGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTGGATGCAAAAGAATTGAATGTACA 2342

Qy   1025 nTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIl 1045
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2343 GTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAGCCTGTGCTCTTCAACTGCAGCAT 2402

Qy   1045 eAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGl 1065
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2403 TGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTGGTGCCATTAGATGAGATCAAAGA 2462

Qy   1065 uAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnTh 1085
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2463 AGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAAGGTCTCCCTGAGAAATACAACAC 2522

Qy   1085 rGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAl 1105
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2523 ACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCTATTGC 2582

Qy   1105 aArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAs 1125
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2583 AAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGA 2642

Qy   1125 pAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCy 1145
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2643 TAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGATAAAGCCAGGACGGGAAGGACATG 2702

Qy   1145 sLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHi 1165
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2703 CCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGGTTCTGCA 2762

Qy   1165 sAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTy 1185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2763 CAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAGCTCCTGAGAAATCGAGACATATA 2822

Qy   1185 rPheLysLeuValAsnAlaGlnSerValGln 1195
          |||||||||||||||||||||||||||||||
Db   2823 TTTTAAGTTAGTGAATGCACAGTCAGTGCAG 2853
```

```
RESULT 5
US-09-866-866A-1
; Sequence 1, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
;  APPLICANT: Sorrentino, Brian
;  APPLICANT:  Schuetz, John
;  TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
;  FILE REFERENCE: 1340-1-021CIP2
;  CURRENT APPLICATION NUMBER: US/09/866,866A
;  CURRENT FILING DATE:  2001-08-30
;  PRIOR APPLICATION NUMBER: 09/584,586
;  PRIOR FILING DATE: 2000-05-31
;  PRIOR APPLICATION NUMBER: PCT/US99/11825
;  PRIOR FILING DATE: 1999-05-27
;  PRIOR APPLICATION NUMBER: 60/086,988
;  PRIOR FILING DATE: 1998-05-28
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 3860
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-866A-1

Alignment Scores:
Pred. No.:              7.94e-311     Length:         3860
Score:                  3554.00       Matches:        682
Percent Similarity:     75.63%        Conservative:   246
Best Local Similarity:  55.58%        Mismatches:     263
Query Match:            59.12%        Indels:         36
DB:                     10            Gaps:           9

US-09-873-409-6 (1-1195) x US-09-866-866A-1 (1-3860)
```

```
Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    151 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210

Qy     21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
Db    211 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 270

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db    271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    331 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 390

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    451 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 510

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    511 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGGTATTGGTGACAAAATT 570

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 630

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    631 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    691 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db    811 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 870

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db    871 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 930

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db    931 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 984

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db    985 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1044

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db   1045 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db   1105 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1164

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db   1165 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1224

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   1225 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1284

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   1285 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1344

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   1345 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1404

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   1405 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1464

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1465 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1524

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1525 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1584

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   1585 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1644

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   1645 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1704

Qy    501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db   1705 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 1764

Qy    521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db   1765 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 1824

Qy    541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db   1825 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 1884

Qy    561 Gln------------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
Db   1885 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 1944

Qy    575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 589
Db   1945 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2004

Qy    590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
Db   2005 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2064

Qy    608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627
Db   2065 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2124

Qy    628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
Db   2125 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2184

Qy    648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
Db   2185 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2244

Qy    667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
Db   2245 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTGGGAATTATTTCTTTTATT 2304

Qy    687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
```

```
              ::::::|||:::||||||        :::|||:::|||||||||||||||||     ||||||
Db     2305 ACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2364

Qy      707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
            |||::::::   |||::::::||||||  ||||||::::::|||||||||:::    :::
Db     2365 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2424

Qy      727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
            |||:::||||||   |||||||||   ||||||   |||   ||||||::::::||||||
Db     2425 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2484

Qy      747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
               |||||||||:::   |||:::|||||||||      |||:::|||      :::|||
Db     2485 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2544

Qy      767 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
            ||||||||||||||||||||||::::::|||   |||:::|||:::|||   |||::::::
Db     2545 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 2604

Qy      787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
            |||:::   |||::::::|||         ::::::|||   |||   |||||||||:::
Db     2605 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 2664

Qy      807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
            ||||||:::   ||||||||||||||||||||||||:::||||||   ||||||:::|||
Db     2665 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 2724

Qy      827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
            |||||||||:::|||:::   ||||||   ||||||   :::   ||||||      :::
Db     2725 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2784

Qy      847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
            ||||||:::   :::||||||   |||   |||      ::::::|||:::   |||
Db     2785 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 2844

Qy      867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
            :::||||||:::||||||         ||||||||||||||||||||||:::
Db     2845 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 2904

Qy      887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
            |||:::   |||   :::   :::||||||:::|||:::   :::||||||||||||:::
Db     2905 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 2964

Qy      907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
            |||:::               ||||||:::|||:::|||||||||      |||||||||:::
Db     2965 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3024

Qy      927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
                  ::::::||||||   |||   |||||||||   |||   ||||||      |||
Db     3025 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3084

Qy      947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
            :::|||   |||||||||:::   |||   ||||||   |||   ||||||   ||||||
Db     3085 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3144

Qy      967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
            |||:::   :::|||:::||||||||||||      :::::::::|||:::|||:::|||
Db     3145 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3204

Qy      987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
            ||||||||||||||||||||||||||||||   |||||||||||||||:::|||   ||||||
Db     3205 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3264

Qy     1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
            |||:::   |||:::||||||   ||||||   :::   |||   ||||||||||||||||||
Db     3265 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3324

Qy     1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
            ||||||:::   :::   ||||||   ||||||||:::||||||:::||||||||||||
Db     3325 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3384

Qy     1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
            |||||||||||||||||||||||||||||||||||         :::||||||      |||
```

```
Db     3385 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3444

Qy     1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
            |||         ||||||||||||:::|||||||||      ||||||      ||||||:::|||:::
Db     3445 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3504

Qy     1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
            ||||||      ||||||      ||||||||||||||||||||||||:::||||||||||||
Db     3505 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3564

Qy     1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
            ||||||:::::::::|||   |||||||||||||||||||||||||||||||||
Db     3565 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 3624

Qy     1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
            :::||||||||||||||||||   ||||||||||||||||||   .||||||||||||:::
Db     3625 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 3684

Qy     1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
            |||:::   ||||||||||||   ||||||||||||||||||||||||||||         |||
Db     3685 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 3744

Qy     1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
            |||::::::||||||   ||||||||||||:::||||||         :::   |||||||||
Db     3745 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 3804

Qy     1187 LysLeuValAsnAlaGlnSer 1193
               :::|||:::   |||:::
Db     3805 TCAATGGTCAGTGTCCAGGCT 3825
```

RESULT 6

```
US-10-044-671-1
; Sequence 1, Application US/10044671
; Patent No. US20020177147A1
; GENERAL INFORMATION:
;  APPLICANT: Washington State University Research Foundation
;  APPLICANT:  Mealey, Katrina
;  APPLICANT:  Bentjen, Steven
;  TITLE OF INVENTION: MDR1 VARIANTS AND METHODS FOR THEIR USE
;  FILE REFERENCE: 4630-61733
;  CURRENT APPLICATION NUMBER: US/10/044,671
;  CURRENT FILING DATE:  2002-01-10
;  PRIOR APPLICATION NUMBER: US 60/261,578
;  PRIOR FILING DATE: 2001-01-12
;  PRIOR APPLICATION NUMBER: US 60/314,829
;  PRIOR FILING DATE: 2001-08-24
;  NUMBER OF SEQ ID NOS: 10
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 4317
;   TYPE: DNA
;   ORGANISM: Canis familiaris
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (70)..(3912)
;   OTHER INFORMATION:
US-10-044-671-1

Alignment Scores:
Pred. No.:               2.65e-310     Length:        4317
Score:                   3549.00       Matches:       680
Percent Similarity:      74.65%        Conservative:  239
Best Local Similarity:   55.24%        Mismatches:    270
Query Match:             59.03%        Indels:        42
DB:                      9             Gaps:          8

US-09-873-409-6 (1-1195) x US-10-044-671-1 (1-4317)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            |||::::::|||    :::|||::::::::::::||||||      ||||||||||||   |||
Db      217 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 276

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
```

```
Db    277 GTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 336

Qy     35 --------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44

Db    337 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 396

Qy     45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59

Db    397 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 456

Qy     60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79

Db    457 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 516

Qy     80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99

Db    517 AAACAATTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 576

Qy    100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118

Db    577 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 636

Qy    119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138

Db    637 AAAGTTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 696

Qy    139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158

Db    697 ACACCTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 756

Qy    159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178

Db    757 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 816

Qy    179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198

Db    817 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 876

Qy    199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218

Db    877 GGACAAAAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGG 936

Qy    219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238

Db    937 ATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCA 996

Qy    239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258

Db    997 TCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TAT 1050

Qy    259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278

Db   1051 ACTATTGGACAGGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGA 1110

Qy    279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298

Db   1111 CAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTC 1170

Qy    299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318

Db   1171 AAGATAATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGAT 1230

Qy    319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338

Db   1231 AATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAA 1290

Qy    339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358

Db   1291 GTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTT 1350

Qy    359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378

Db   1351 GGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCC 1410

Qy    379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398

Db   1411 ACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTT 1470

Qy    399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418

Db   1471 CGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAA 1530

Qy    419 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438

Db   1531 AACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAA 1590

Qy    439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458

Db   1591 GCCAATGCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAG 1650

Qy    459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478

Db   1651 AGAGGGGCCCGGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTT 1710

Qy    479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498

Db   1711 CGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAA 1770

Qy    499 SerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 518

Db   1771 GCAGTGGTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCT 1830

Qy    519 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 538

Db   1831 CATCGTTTGTCTACAGTTCGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATT 1890

Qy    539 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 558

Db   1891 GTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTC 1950

Qy    559 MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 578

Db   1951 ACAATGCAG---ACAAGAGGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAA 2007

Qy    579 ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--- 597

Db   2008 AGTGAAAGTGATGCCTTGGAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAA 2064

Qy    598 -----------------------------------------AspLysAlaGluGluSer 603

Db   2065 AGAAGATCAACTCGCAGGAGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACA 2124

Qy    604 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 623

Db   2125 AAAGAGGACTTGAATGAGAATGTTCCTTCAGTTTCCTTCTGGAGGATTCTGAAGCTGAAC 2184

Qy    624 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 643

Db   2185 TCAACTGAATGGCCTTATTTTGTGGTTGGTATATTTTGTGCTATTATAAACGGAGGCCTG 2244

Qy    644 HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAsp 662

Db   2245 CAACCAGCATTTTCAATAATATTTTCAAGGATTATAGGGATCTTTACCCGAGATGAGGAT 2304

Qy    663 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 682

Db   2305 CCTGAAACAAAACGACAGAATAGTAACATGTTTTCTGTATTGTTTCTAGTCCTTGGAATT 2364

Qy    683 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 702

Db   2365 ATTTCTTTTATTACATTTTTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTC 2424

Qy    703 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 722

Db   2425 ACTAAGCGGCTTCGATACATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTT 2484
```

```
Qy    723 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 742
Db   2485 GATGACCCTAAAAACACCACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAA 2544

Qy    743 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 762
Db   2545 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGG 2604

Qy    763 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 782
Db   2605 ACAGGCATTATTATATCCTTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATT 2664

Qy    783 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 802
Db   2665 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 2724

Qy    803 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 822
Db   2725 AAAGATAAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTC 2784

Qy    823 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 842
Db   2785 CGAACTGTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTG 2844

Qy    843 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 862
Db   2845 CAAGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATC 2904

Qy    863 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 882
Db   2905 ACCCAGGCAATGATGTATTTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTG 2964

Qy    883 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 902
Db   2965 GCAAATGAGTTCATGAACTTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGT 3024

Qy    903 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 922
Db   3025 GCCATGGCAGTGGGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCA 3084

Qy    923 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 942
Db   3085 GCAGCCCACGTCATCATGATCATTGAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCAC 3144

Qy    943 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 962
Db   3145 GGCCTCAAGCCAAATACGTTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTAT 3204

Qy    963 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 982
Db   3205 CCCACTCGACCAGACATCCCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAG 3264

Qy    983 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 1002
Db   3265 ACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAG 3324

Qy   1003 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 1022
Db   3325 CGCTTCTATGACCCCTTGGCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTG 3384

Qy   1023 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 1042
Db   3385 AATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGAC 3444

Qy   1043 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 1062
Db   3445 TGCAGCATTGCCGAGAACATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAG 3504

Qy   1063 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 1082
Db   3505 ATTATGCAGGCAGCCAAGGAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAA 3564

Qy   1083 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1102
Db   3565 TACAACACCAGAGTAGGAGACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATT 3624

Qy   1103 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 1122
Db   3625 GCCATAGCTCGCGCTCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCTACATCA 3684

Qy   1123 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1142
Db   3685 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 3744

Qy   1143 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1162
Db   3745 CGCACCTGCATTGTGATCGCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTG 3804

Qy   1163 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1182
Db   3805 GTGTTTCAGAATGGCAAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGGCTCAGAAA 3864

Qy   1183 AspIleTyrPheLysLeuValAsnAlaGlnSer 1193
Db   3865 GGCATCTATTTTTCCATGATCAGTGTCCAGGCT 3897
```

RESULT 7
US-09-866-866A-3
; Sequence 3, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 3860
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-866-866A-3

Alignment Scores:

Pred. No.:	5.16e-310	Length:	3860
Score:	3545.00	Matches:	681
Percent Similarity:	75.55%	Conservative:	246
Best Local Similarity:	55.50%	Mismatches:	264
Query Match:	58.97%	Indels:	36
DB:	10	Gaps:	9

US-09-873-409-6 (1-1195) x US-09-866-866A-3 (1-3860)

```
Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    151 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210

Qy     21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
Db    211 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 270

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db    271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
```

```
Db     331 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 390

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db     391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db     451 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 510

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db     511 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 570

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db     571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 630

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db     631 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db     691 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db     751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810

Qy     201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db     811 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 870

Qy     221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db     871 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 930

Qy     241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db     931 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 984

Qy     261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db     985 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1044

Qy     281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db    1045 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104

Qy     301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db    1105 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1164

Qy     321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db    1165 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1224

Qy     341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db    1225 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1284

Qy     361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db    1285 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1344

Qy     381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db    1345 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1404

Qy     401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db    1405 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1464

Qy     421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db    1465 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1524

Qy     441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db    1525 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1584

Qy     461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db    1585 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1644

Qy     481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db    1645 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1704

Qy     501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db    1705 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 1764

Qy     521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db    1765 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 1824

Qy     541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db    1825 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 1884

Qy     561 Gln-----------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
Db    1885 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 1944

Qy     575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 589
Db    1945 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2004

Qy     590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
Db    2005 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2064

Qy     608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627
Db    2065 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2124

Qy     628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
Db    2125 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2184

Qy     648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
Db    2185 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2244

Qy     667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
Db    2245 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2304

Qy     687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
Db    2305 ACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2364

Qy     707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
Db    2365 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2424

Qy     727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
Db    2425 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2484

Qy     747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
Db    2485 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2544
```

```
Qy    767 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
          ||||||||||||||||||||||:::::|||  |||:::|||:::|||  |||::::::
Db   2545 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 2604

Qy    787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
          |||:::   |||:::::|||          ::::::|||  |||  |||||||||:::
Db   2605 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 2664

Qy    807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
          ||||||:::   ||||||||||||||||||||||||:::||||||   ||||||:::|||
Db   2665 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 2724

Qy    827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
          |||||||||:::|||:::   ||||||   ||||||   :::   ||||||      :::
Db   2725 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2784

Qy    847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
          ||||||:::   :::||||||   |||   |||      ::::::|||:::   |||
Db   2785 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 2844

Qy    867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
          :::||||||:::||||||      ||||||||||||||||||||||||:::
Db   2845 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 2904

Qy    887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
          |||:::   |||   :::   :::||||||:::|||:::   :::||||||||||||:::
Db   2905 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 2964

Qy    907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
          |||:::            ||||||:::|||:::|||||||||      |||||||||:::
Db   2965 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3024

Qy    927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
                ::::::||||||   |||   |||||||||   |||   ||||||      |||
Db   3025 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3084

Qy    947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
          :::|||   ||||||||||:::   ||||||   ||||||   |||   ||||||   ||||||
Db   3085 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3144

Qy    967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
          |||:::   :::|||:::|||||||||||||   :::::::::|||:::|||:::|||
Db   3145 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3204

Qy    987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
          ||||||||||||||||||||||||||||||   ||||||||||||:::|||   ||||||
Db   3205 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3264

Qy   1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
          |||:::   |||:::||||||   ||||||   :::   |||   ||||||||||||||
Db   3265 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3324

Qy   1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
          ||||||:::   :::   ||||||   ||||||||||:::||||||:::||||||||||||
Db   3325 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3384

Qy   1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
          |||||||||||||||||||||||||||||||||||||         :::||||||      |||
Db   3385 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3444

Qy   1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
          |||   ||||||||||||:::||||||||||   ||||||   ||||||:::|||:::
Db   3445 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3504

Qy   1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
          ||||||   ||||||   ||||||||||||||||||||||||:::||||||||||||
Db   3505 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3564

Qy   1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
          ||||||::::::::|||   ||||||||||||||||||||||||||||||||||||
Db   3565 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 3624

Qy   1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
          :::||||||||||||||||||   ||||||||||||||||   |||||||||||:::
Db   3625 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 3684

Qy   1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
          |||:::   ||||||||||||   ||||||||||||||||||||||||||      |||
Db   3685 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 3744

Qy   1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
          |||::::::||||||   ||||||||||||:::||||||      :::   |||||||||
Db   3745 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 3804

Qy   1187 LysLeuValAsnAlaGlnSer 1193
             :::|||:::   |||:::
Db   3805 TCAATGGTCAGTGTCCAGGCT 3825
```

RESULT 8
US-09-306-417-1
; Sequence 1, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 8630
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: proviral
;   OTHER INFORMATION: plasmid DNA
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(160)
;   OTHER INFORMATION: plasmid backbone (pUC)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (161)..(677)
;   OTHER INFORMATION: 5'-LTR
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (532)..(1219)
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: (1220)..(5062)
;   OTHER INFORMATION: m4 mdr-1 cDNA
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5215)..(5774)
;   OTHER INFORMATION: 3'-LTR
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5775)..(8630)
;   OTHER INFORMATION: plasmid backbone (pUC)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(8630)
;   OTHER INFORMATION: retroviral expression vector SFbeta71m4
US-09-306-417-1

Alignment Scores:

Pred. No.:	1.67e-309	Length:	8630
Score:	3545.00	Matches:	681
Percent Similarity:	75.55%	Conservative:	246
Best Local Similarity:	55.50%	Mismatches:	264
Query Match:	58.97%	Indels:	36
DB:	10	Gaps:	9

```
US-09-873-409-6 (1-1195) x US-09-306-417-1 (1-8630)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    1370 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 1429

Qy      21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db    1430 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 1489

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db    1490 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 1549

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    1550 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 1609

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    1610 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 1669

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    1670 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 1729

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    1730 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 1789

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    1790 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1849

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    1850 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1909

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    1910 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1969

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    1970 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 2029

Qy     201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db    2030 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 2089

Qy     221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db    2090 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 2149

Qy     241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db    2150 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 2203

Qy     261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db    2204 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTAGTGTTGGACAGGCA 2263

Qy     281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db    2264 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 2323

Qy     301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db    2324 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 2383

Qy     321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db    2384 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 2443

Qy     341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db    2444 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 2503

Qy     361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db    2504 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 2563

Qy     381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db    2564 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 2623

Qy     401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db    2624 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 2683

Qy     421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db    2684 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 2743

Qy     441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db    2744 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2803

Qy     461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db    2804 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2863

Qy     481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db    2864 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2923

Qy     501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db    2924 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 2983

Qy     521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db    2984 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 3043

Qy     541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db    3044 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 3103

Qy     561 Gln-----------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
Db    3104 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 3163

Qy     575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 589
Db    3164 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 3223

Qy     590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
Db    3224 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 3283

Qy     608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627
Db    3284 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 3343

Qy     628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
Db    3344 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 3403

Qy     648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
Db    3404 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 3463

Qy     667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
Db    3464 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 3523
```

```
Qy    687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
Db   3524 ACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 3583

Qy    707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
Db   3584 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 3643

Qy    727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
Db   3644 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 3703

Qy    747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
Db   3704 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 3763

Qy    767 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
Db   3764 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 3823

Qy    787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
Db   3824 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 3883

Qy    807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
Db   3884 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 3943

Qy    827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
Db   3944 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 4003

Qy    847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
Db   4004 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 4063

Qy    867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
Db   4064 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 4123

Qy    887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
Db   4124 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 4183

Qy    907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
Db   4184 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 4243

Qy    927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
Db   4244 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 4303

Qy    947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
Db   4304 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 4363

Qy    967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
Db   4364 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 4423

Qy    987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
Db   4424 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 4483

Qy   1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
Db   4484 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATTAAGCGACTGAATGTTCAGTGG 4543

Qy   1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
Db   4544 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 4603

Qy   1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
Db   4604 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 4663

Qy   1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
Db   4664 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 4723

Qy   1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
Db   4724 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 4783

Qy   1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
Db   4784 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 4843

Qy   1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
Db   4844 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 4903

Qy   1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
Db   4904 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 4963

Qy   1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
Db   4964 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 5023

Qy   1187 LysLeuValAsnAlaGlnSer 1193
Db   5024 TCAATGGTCAGTGTCCAGGCT 5044
```

```
RESULT 9
US-09-306-417-2
; Sequence 2, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
;  APPLICANT: Heinrich-Pette-Institut
;  TITLE OF INVENTION: Retroviral Gene Transfer Vectors
;  FILE REFERENCE: P50491
;  CURRENT APPLICATION NUMBER: US/09/306,417
;  CURRENT FILING DATE: 1999-05-06
;  EARLIER APPLICATION NUMBER: DE 198 22 115
;  EARLIER FILING DATE: 1998-05-08
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 8630
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: proviral
;   OTHER INFORMATION: plasmid DNA
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(8630)
;   OTHER INFORMATION: retroviral expression vector SFbeta91mSA1
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(160)
;   OTHER INFORMATION: plasmid backbone (pUC)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (161)..(677)
;   OTHER INFORMATION: 5'-LTR
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (532)..(1219)
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: (1220)..(5062)
;   OTHER INFORMATION: mSA1 mdr1 cDNA
;   FEATURE:
;   NAME/KEY: misc_feature
```

```
;    LOCATION: (5215)..(5774)
;    OTHER INFORMATION: 3'-LTR
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (5775)..(8630)
;    OTHER INFORMATION: plasmid backbone (pUC)
US-09-306-417-2

Alignment Scores:
Pred. No.:               1.67e-309    Length:         8630
Score:                   3545.00      Matches:        681
Percent Similarity:      75.55%       Conservative:   246
Best Local Similarity:   55.50%       Mismatches:     264
Query Match:             58.97%       Indels:         36
DB:                      10           Gaps:           9

US-09-873-409-6 (1-1195) x US-09-306-417-2 (1-8630)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db   1370 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 1429

Qy     21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
Db   1430 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 1489

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db   1490 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 1549

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db   1550 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 1609

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db   1610 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 1669

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db   1670 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 1729

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db   1730 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 1789

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db   1790 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1849

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db   1850 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1909

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db   1910 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1969

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db   1970 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 2029

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db   2030 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 2089

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db   2090 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 2149

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db   2150 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 2203

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db   2204 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 2263

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db   2264 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 2323

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db   2324 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 2383

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db   2384 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 2443

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   2444 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 2503

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   2504 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 2563

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   2564 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 2623

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   2624 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 2683

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   2684 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 2743

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   2744 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2803

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   2804 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2863

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   2864 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2923

Qy    501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db   2924 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 2983

Qy    521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db   2984 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 3043

Qy    541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db   3044 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 3103

Qy    561 Gln------------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
Db   3104 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 3163

Qy    575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 589
Db   3164 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 3223

Qy    590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
Db   3224 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 3283
```

```
Qy    608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627
Db   3284 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 3343

Qy    628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
Db   3344 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 3403

Qy    648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
Db   3404 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 3463

Qy    667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
Db   3464 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 3523

Qy    687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
Db   3524 ACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 3583

Qy    707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
Db   3584 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 3643

Qy    727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
Db   3644 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 3703

Qy    747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
Db   3704 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 3763

Qy    767 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
Db   3764 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 3823

Qy    787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
Db   3824 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 3883

Qy    807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
Db   3884 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 3943

Qy    827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
Db   3944 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 4003

Qy    847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
Db   4004 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 4063

Qy    867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
Db   4064 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 4123

Qy    887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
Db   4124 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 4183

Qy    907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
Db   4184 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 4243

Qy    927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
Db   4244 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 4303

Qy    947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
Db   4304 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 4363

Qy    967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
Db   4364 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 4423

Qy    987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
Db   4424 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 4483

Qy   1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
Db   4484 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 4543

Qy   1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
Db   4544 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 4603

Qy   1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
Db   4604 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 4663

Qy   1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
Db   4664 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 4723

Qy   1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
Db   4724 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 4783

Qy   1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
Db   4784 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 4843

Qy   1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
Db   4844 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 4903

Qy   1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
Db   4904 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 4963

Qy   1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
Db   4964 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 5023

Qy   1187 LysLeuValAsnAlaGlnSer 1193
Db   5024 TCAATGGTCAGTGTCCAGGCT 5044
```

RESULT 10
US-09-769-097-3
; Sequence 3, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: ENCODING RAT MDR1A AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP-50009-C2
; CURRENT APPLICATION NUMBER: US/09/769,097
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 4425
; TYPE: DNA

```
;   ORGANISM: RATTUS RATTUS
US-09-769-097-3

Alignment Scores:
Pred. No.:              1.6e-309      Length:         4425
Score:                  3540.50       Matches:        680
Percent Similarity:     75.53%        Conservative:   243
Best Local Similarity:  55.65%        Mismatches:     266
Query Match:            58.89%        Indels:         33
DB:                     10            Gaps:           9

US-09-873-409-6 (1-1195) x US-09-769-097-3 (1-4425)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db     499 ATGCTGCTGGGAACTCTGGCGGCCATTATCCATGGAATTGCGCTCCCACTTATGATGCTG 558

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db     559 GTCTTTGGAGACATGACAGATAGCTTTGCAAAT------GTAGGAAACAACCGTAGTATG 612

Qy      41 SerPhePheArgLeuThr------------------------------------------ 46
Db     613 AGTTTCTACAATGCTACAGACATATATGCCAAGCTGGAGGACGAAATGACCACGTACGCC 672

Qy      47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
Db     673 TACTATTACACGGGCATTGGTGCCGGTGTGCTCATCGTTGCCTACATCCAGGTTTCCACT 732

Qy      67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
Db     733 TGGTGCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAAGTTTTTCCATGCCATC 792

Qy      87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
Db     793 ATGAATCAGGAGATAGGCTGGTTTGACGTGCATGACGTTGGGGAGCTCAACACCCGGCTC 852

Qy     107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
Db     853 ACAGATGACGTCTCCAAAATTAATGAAGGAATTGGTGACAAAATTGGAATGTTCTTTCAG 912

Qy     126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
Db     913 GCAATGGCAACATTTTTTGGTGGTTTTATAATAGGATTTACTCGCGGCTGGAAGCTAACT 972

Qy     146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
Db     973 CTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTGTCAGCTGGTATTTGGGCAAAGATA 1032

Qy     166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
Db    1033 TTGTCTTCATTTACTGATAAGGAACTCCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAA 1092

Qy     186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
Db    1093 GAAGTCTTAGCAGCCATCAGAACTGTGATTGCCTTTGGAGGACAAAAGAAGGAACTTGAA 1152

Qy     206 ArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSer 225
Db    1153 AGGTACAATAACAATTTGGAAGAAGCTAAAAGGCTTGGGATAAAGAAAGCTATCACGGCC 1212

Qy     226 LysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrp 245
Db    1213 AACATTTCCATGGGTGCAGCTTTTCTGCTTATCTATGCATCATATGCTCTGGCATTCTGG 1272

Qy     246 TyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAla 265
Db    1273 TATGGGACTTCCTTGGTCATCTCAAAAGAA------TACACTATTGGACAAGTGCTCACT 1326

Qy     266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285
Db    1327 GTCTTTTTTTCTGTATTAATTGGAGCATTCAGTGTTGGGCAGGCATCTCCAAATATTGAA 1386

Qy     286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305
Db    1387 GCCTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTCAGTATAATTGATAATAAGCCC 1446

Qy     306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
Db    1447 AGTATAGACAGCTTCTCAAAGAGTGGGCACAAACCCGACAACATACAAGGAAATTTGGAA 1506

Qy     326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
Db    1507 TTCAAAAATATTCACTTCAGTTACCCGTCTCGAAAAGACGTTCAGATCTTGAAGGGCCTC 1566

Qy     346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
Db    1567 AACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTGGTTGGCAACAGTGGCTGTGGGAAA 1626

Qy     366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
Db    1627 AGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCATAGAGGGCGAGGTCAGTATC 1686

Qy     386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
Db    1687 GACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTGGTG 1746

Qy     406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
Db    1747 AGTCAGGAACCCGTGCTGTTTGCCACCACAATTGCCGAAAACATTCGCTATGGCCGAGAA 1806

Qy     426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
Db    1807 AACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAATGCCTATGATTTCATC 1866

Qy     446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465
Db    1867 ATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGG 1926

Qy     466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
Db    1927 GGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAACCCCAAGATCCTTTTG 1986

Qy     486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeu 505
Db    1987 TTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTGGTTCAGGCCGCTCTG 2046

Qy     506 GluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArg 525
Db    2047 GATAAGGCTAGAGAAGGCCGGACCACCATTGTGATAGCTCACCGCTTGTCTACAGTTCGC 2106

Qy     526 SerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAla 545
Db    2107 AATGCTGAYGTCATTGCTGGTTTTGATGGTGGTGTCATTGTGGAGCAAGGAAATCATGAT 2166

Qy     546 GluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln------------ 561
Db    2167 GAGCTCATGAGAGAGAAAGGAATTTACTTCAAACTTGTCATGACTCAGACAGCAGGAAAT 2226

Qy     562 AspIleLysLysAlaAspGluGlnMetGluSer------------------MetThrTyr 575
Db    2227 GAAATTGAATTAGGAAATGAAGCTTGTGAATCTAAAGAYGGAATTGATAATGTGGACATG 2286

Qy     576 SerThrGluArgLysThrAsnSerLeu------ProLeuHisSerValLysSerIleLys 593
Db    2287 TCTTCAAAAGATTCRGGATCCAGTCTAATAAGAAGAAGATCAACTCGCAAAAGCATCCGT 2346

Qy     594 SerAspPheIleAspLysAlaGluGluSerThr---GlnSerLysGluIleSerLeuPro 612
Db    2347 GGGCCACATGATCAAGACGGGGAACTTAGCACCAAAGAGGCTCTGGATGACGACGTACCT 2406

Qy     613 GluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeu 632
Db    2407 CCAGCTTCCTTTTGGCGGATCCTGAAGTTGAATTCAACTGAATGGCCTTATTTTGTGGTT 2466

Qy     633 GlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAla 652
```

```
Db   2467 GGTGTATTTTGTGCCATAATAAATGGAGGCTTGCAACCAGCATTCTCCATAATATTTTCA 2526

Qy    653 LysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeu---LysHisAspAlaGlu 671
Db   2527 AAGGTTGTAGGGGTTTTTACAAAAAATGACACCCCTGAAATCCAGCGGCAGAACAGCAAC 2586

Qy    672 IleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGln 691
Db   2587 TTGTTTTCTTTATTGTTCCTGATCCTTGGGATCATCTCTTTCATTACGTTTTTCCTTCAG 2646

Qy    692 GlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPhe 711
Db   2647 GGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGACTCCGATACATGGTCTTC 2706

Qy    712 LysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGly 731
Db   2707 AAATCCATGCTGAGACAGGACATAAGCTGGTTTGATGACCCTAAAAACACCACAGGAGCG 2766

Qy    732 LeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIle 751
Db   2767 CTGACCACCAGGCTTGCCAATGACGCTGCTCAAGTGAAAGGGGCTACAGGGTCTAGGCTT 2826

Qy    752 GlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyr 771
Db   2827 GCTGTTATTACCCAGAACATAGCAAATCTTGGGACAGGCATCATCATATCCCTGATCTAC 2886

Qy    772 GlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMet 791
Db   2887 GGCTGGCAATTGACACTTTTACTCCTAGCAATTGTTCCCATCATTGCTATAGCAGGAGTG 2946

Qy    792 IleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAla 811
Db   2947 GTTGAAATGAAAATGTTGTCTGGACAAGCGCTGAAAGATAAGAAGGAACTAGAAGGTTCT 3006

Qy    812 GlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGlu 831
Db   3007 GGGAAGATCGCTACAGAAGCAATTGAAAACTTTCGCACTGTCGTCTCTTTGACTCGGGAG 3066

Qy    832 LysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLys 851
Db   3067 CAGAAGTTTGAAACTATGTATGCCCAGAGCTTGCAGATACCATACAGAAATGCTTTGAAG 3126

Qy    852 LysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyr 871
Db   3127 AAAGCGCACGTCTTTGGGATCACTTTCTCCTTCACCCAGGCCATGATGTATTTCTCCTAT 3186

Qy    872 AlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGly 891
Db   3187 GCTGCTTGTTTCCGGTTTGATGCCTACTTGGTGGCACGAGAACTCATGACATTTGAAAAT 3246

Qy    892 MetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuVal 911
Db   3247 GTTCTGTTAGTATTCTCAGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAGGTCAGTTCA 3306

Qy    912 LeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGlu 931
Db   3307 TTCGCTCCTGACTACGCGAAAGCCAAAGTCTCGGCATCCCACATCATCAGGATCATTGAG 3366

Qy    932 LysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGly 951
Db   3367 AAAATCCCTGAGATTGACAGCTACAGCACGGAGGGCTTGAAGCCTAATATGTTGGAAGGA 3426

Qy    952 AsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeu 971
Db   3427 AATGTGAAATTTAATGGAGTCATGTTCAACTATCCCACCCGACCCAACATCCCAGTGCTT 3486

Qy    972 ArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGly 991
Db   3487 CAGGGGCTGAGCCTAGAGGTGAAGAAAGGGCAGACGCTGGCCCTCGTGGGCAGCAGTGGC 3546

Qy    992 CysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGln 1011

Db   3547 TGCGGGAAGAGTACAGTGGTCCAGCTGCTTGAGCGCTTCTATGACCCCATGGCCGGAACA 3606

Qy   1012 ValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIle 1031
Db   3607 GTGTTTCTAGATGGCAAAGAAATAAAGCAACTCAATGTCCAGTGGCTCCGCGCCCACCTG 3666

Qy   1032 AlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyr 1051
Db   3667 GGCATTGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATCGCCGAGAACATTGCCTAC 3726

Qy   1052 GlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsn 1071
Db   3727 GGAGACAACAGCCGTGTCGTGTCTCATAAGGAGATCGTGAAGGCAGCCAAGGAGGCCAAC 3786

Qy   1072 IleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGly 1091
Db   3787 ATCCACCAGTTCATCGACTCACTGCCTGAGAAATACAACACCAGAGTGGGAGACAAAGGG 3846

Qy   1092 AlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLys 1111
Db   3847 ACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATCGCCATCGCGCGCGCCCTCGTCAGACAG 3906

Qy   1112 ProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysVal 1131
Db   3907 CCTCACATCTTACTTCTGGATGAAGCGACATCAGCTCTGGATACGGAGAGTGAAAAGGTC 3966

Qy   1132 ValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArg 1151
Db   3967 GTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACCTGCATTGTGATCGCGCACCGC 4026

Qy   1152 LeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGlu 1171
Db   4027 CTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATTCAGAACGGCCAGGTCAAGGAG 4086

Qy   1172 GlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 1191
Db   4087 CACGGCACCCACCAGCAGCTGCTGGCCCAGAAAGGCATCTATTTCTCGATGGTCAGTGTG 4146

Qy   1192 GlnSer 1193
Db   4147 CAGGCT 4152
```

RESULT 11
US-09-769-097-1
; Sequence 1, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: ENCODING RAT MDR1A AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP-50009-C2
; CURRENT APPLICATION NUMBER: US/09/769,097
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4369
; TYPE: DNA
; ORGANISM: RATTUS RATTUS
US-09-769-097-1

Alignment Scores:
Pred. No.: 2.39e-309 Length: 4369

```
Score:                      3538.50      Matches:        680
Percent Similarity:         75.53%       Conservative:   243
Best Local Similarity:      55.65%       Mismatches:     266
Query Match:                58.86%       Indels:         33
DB:                         10           Gaps:           9

US-09-873-409-6 (1-1195) x US-09-769-097-1 (1-4369)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    499 ATGCTGCTGGGAACTCTGGCGGCCATTATCCATGGAATTGCGCTCCCACTTATGATGCTG 558

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db    559 GTCTTTGGAGACATGACAGATAGCTTTGCAAAT------GTAGGAAACAACCGTAGTATG 612

Qy     41 SerPhePheArgLeuThr---------------------------------------- 46
Db    613 AGTTTCTACAATGCTACAGACATATATGCCAAGCTGGAGGACGAAATGGCCACGTACGCC 672

Qy     47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
Db    673 TACTATTACACGGGCAATGGTGCCGGTGTGCTCATCGTTGCCTACATCCAGGTTTCACTT 732

Qy     67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
Db    733 TGGTGCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAAGTTTTTCCATGCCATC 792

Qy     87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
Db    793 ATGAATCAGGAGATAGGCTGGTTTGACGTGCATGACGTTGGGGAGCTCAACACCCGGCTC 852

Qy    107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
Db    853 ACAGATGACGTCTCCAAAATTAATGAAGGAATTGGTGACAAAATTGGAATGTTCTTTCAG 912

Qy    126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
Db    913 GCAATGGCAACATTTTTTGGTGGTTTTATAATAGGATTTACTCGCGGCTGGAAGCTAACT 972

Qy    146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
Db    973 CTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTGTCAGCTGGTATTTGGGCAAAGATA 1032

Qy    166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
Db   1033 TTGTCTTCATTTACTGATAAGGAACTCCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAA 1092

Qy    186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
Db   1093 GAAGTCTTAGCAGCCATCAGAACTGTGATTGCCTTTGGAGGACAAAAGAAGGAACTTGAA 1152

Qy    206 ArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSer 225
Db   1153 AGGTACAATAACAATTTGGAAGAAGCTAAAAGGCTTGGGATAAAGAAAGCTATCACGGCC 1212

Qy    226 LysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrp 245
Db   1213 AACATTTCCATGGGTGCAGCTTTTCTGCTTATCTATGCATCATATGCTCTGGCATTCTGG 1272

Qy    246 TyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAla 265
Db   1273 TATGGGACTTCCTTGGTCATCTCAAAAGAA------TACACTATTGGACAAGTGCTCACT 1326

Qy    266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285
Db   1327 GTCTTTTTTTCTGTATTAATTGGAGCATTCAGTGTTGGCAGGCATCTCCAAATATTGAA 1386

Qy    286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305
Db   1387 GCCTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTCAGTATAATTGATAATAAGCCC 1446

Qy    306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
Db   1447 AGTATAGACAGCTTCTCAAAGAGTGGGCACAAACCCGACAACATACAAGGAAATTTGGAA 1506

Qy    326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
Db   1507 TTCAAAAATATTCACTTCAGTTACCCGTCTCGAAAAGACGTTCAGATCTTGAAGGGCCTC 1566

Qy    346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
Db   1567 AACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTGGTTGGCAACAGTGGCTGTGGGAAA 1626

Qy    366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
Db   1627 AGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCATAGAGGGCGAGGTCAGTATC 1686

Qy    386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
Db   1687 GACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTGGTG 1746

Qy    406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
Db   1747 AGTCAGGAACCCGTGCTGTTTGCCACCACAATTGCCGAAAACATTCGCTATGGCCGAGAA 1806

Qy    426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
Db   1807 AACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAATGCCTATGATTTCATC 1866

Qy    446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465
Db   1867 ATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGG 1926

Qy    466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
Db   1927 GGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAACCCCAAGATCCTTTTG 1986

Qy    486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeu 505
Db   1987 TTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTGGTTCAGGCCGCTCTG 2046

Qy    506 GluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArg 525
Db   2047 GATAAGGCTAGAGAAGGCCGGACCACCATTGTGATAGCTCACCGCTTGTCTACAGTTCGC 2106

Qy    526 SerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAla 545
Db   2107 AATGCTGACGTCATTGCTGGTTTTGATGGTGGTGTCATTGTGGAGCAAGGAAATCATGAT 2166

Qy    546 GluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln------------ 561
Db   2167 GAGCTCATGAGAGAGAAAGGAATTTACTTCAAACTTGTCATGACTCAGACAGCAGGAAAT 2226

Qy    562 AspIleLysLysAlaAspGluGlnMetGluSer------------------MetThrTyr 575
Db   2227 GAAATTGAATTAGGAAATGAAGCTTGTGAATCTAAAGACGGAATTGATAATGTGGACATG 2286

Qy    576 SerThrGluArgLysThrAsnSerLeu------ProLeuHisSerValLysSerIleLys 593
Db   2287 TCTTCAAAAGATTCGGGATCCAGTCTAATAAGAAGAAGATCAACTCGCAAAAGCATCCGT 2346

Qy    594 SerAspPheIleAspLysAlaGluGluSerThr---GlnSerLysGluIleSerLeuPro 612
Db   2347 GGGCCACATGATCAAGACGGGGAACTTAGCACCAAAGAGGCTCTGGATGACGACGTACCT 2406

Qy    613 GluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeu 632
Db   2407 CCAGCTTCCTTTTGGCGGATCCTGAAGTTGAATTCAACTGAATGGCCTTATTTTGTGGTT 2466

Qy    633 GlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAla 652
Db   2467 GGTGTATTTTGTGCCATAATAAATGGAGGCTTGCAACCAGCATTCTCCATAATATTTTCA 2526

Qy    653 LysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeu---LysHisAspAlaGlu 671
```

```
Db    2527 AAGGTTGTAGGGGTTTTACAAAAAATGACACCCCTGAAATCCAGCGGCAGAACAGCAAC 2586

Qy     672 IleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGln 691
           ::::::|||::::::|||:::||||||||:::|||   |||:::::::::|||:::|||
Db    2587 TTGTTTTCTTTATTGTTCCTGATCCTTGGGATCATCTCTTTCATTACGTTTTTCCTTCAG 2646

Qy     692 GlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPhe 711
           |||      :::|||:::||||||||||||||||||   |||||||||::::::   |||
Db    2647 GGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGACTCCGATACATGGTCTTC 2706

Qy     712 LysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGly 731
           |||:::||||||   ||||||||:::|||||||||:::   :::|||:::||||||
Db    2707 AAATCCATGCTGAGACAGGACATAAGCTGGTTTGATGACCCTAAAAACACCACAGGAGCG 2766

Qy     732 LeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIle 751
           |||||||||   ||||||   |||   ||||||:::::::||||||||||||||||||:::
Db    2767 CTGACCACCAGGCTTGCCAATGACGCTGCTCAAGTGAAAGGGGCTACAGGGTCTAGGCTT 2826

Qy     752 GlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyr 771
              |||:::|||||||||      |||:::|||      :::|||||||||   ||||||
Db    2827 GCTGTTATTACCCAGAACATAGCAAATCTTGGGACAGGCATCATCATATCCCTGATCTAC 2886

Qy     772 GlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMet 791
           ||||||::::::|||   |||:::|||:::|||   |||::::::|||:::   |||:::
Db    2887 GGCTGGCAATTGACACTTTTACTCCTAGCAATTGTTCCCATCATTGCTATAGCAGGAGTG 2946

Qy     792 IleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAla 811
           :::|||         ::::::|||   |||   |||||||||:::||||||:::   :::
Db    2947 GTTGAAATGAAAATGTTGTCTGGACAAGCGCTGAAAGATAAGAAGGAACTAGAAGGTTCT 3006

Qy     812 GlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGlu 831
           ||||||||||||||||||:::||||||   ||||||:::||||||||||||||||||
Db    3007 GGGAAGATCGCTACAGAAGCAATTGAAAACTTTCGCACTGTCGTCTCTTTGACTCGGGAG 3066

Qy     832 LysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLys 851
           :::   ||||||   ||||||   :::   ||||||      :::||||||      |||
Db    3067 CAGAAGTTTGAAACTATGTATGCCCAGAGCTTGCAGATACCATACAGAAATGCTTTGAAG 3126

Qy     852 LysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyr 871
           ||||||   :::   |||      ::::::|||:::   |||   :::||||||:::|||
Db    3127 AAAGCGCACGTCTTTGGGATCACTTTCTCCTTCACCCAGGCCATGATGTATTTCTCCTAT 3186

Qy     872 AlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGly 891
           ||||||   |||||||||   |||||||||:::               ||||||   |||
Db    3187 GCTGCTTGTTTCCGGTTTGATGCCTACTTGGTGGCACGAGAACTCATGACATTTGAAAAT 3246

Qy     892 MetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuVal 911
           :::   :::||||||:::||||||   :::||||||||||||:::||||||:::
Db    3247 GTTCTGTTAGTATTCTCAGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAGGTCAGTTCA 3306

Qy     912 LeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGlu 931
              ||||||:::|||:::|||||||||      |||:::|||:::      ::::::|||
Db    3307 TTCGCTCCTGACTACGCGAAAGCCAAAGTCTCGGCATCCCACATCATCAGGATCATTGAG 3366

Qy     932 LysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGly 951
           |||   |||   |||||||||   |||   ||||||   ||||||:::      ||||||
Db    3367 AAAATCCCTGAGATTGACAGCTACAGCACGGAGGGCTTGAAGCCTAATATGTTGGAAGGA 3426

Qy     952 AsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeu 971
           |||::::::|||      |||   |||   ||||||   ||||||::::::   :::|||
Db    3427 AATGTGAAATTTAATGGAGTCATGTTCAACTATCCCACCCGACCCAACATCCCAGTGCTT 3486

Qy     972 ArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGly 991
           :::|||||||||||||      :::::::::|||:::|||   ||||||||||||||
Db    3487 CAGGGGCTGAGCCTAGAGGTGAAGAAAGGGCAGACGCTGGCCCTCGTGGGCAGCAGTGGC 3546

Qy     992 CysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGln 1011
           |||||||||||||||   ||||||||||||:::|||   |||||||||:::   |||
Db    3547 TGCGGGAAGAGTACAGTGGTCCAGCTGCTTGAGCGCTTCTATGACCCCATGGCCGGAACA 3606

Qy    1012 ValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIle 1031
           |||   ||||||   :::   |||:::||||||||||||||||||||||:::   :::
Db    3607 GTGTTTCTAGATGGCAAAGAAATAAAGCAACTCAATGTCCAGTGGCTCCGCGCCCACCTG 3666

Qy    1032 AlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyr 1051
              ||||||   |||||||||:::||||||:::||||||||||||||||||||||||
Db    3667 GGCATTGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATCGCCGAGAACATTGCCTAC 3726

Qy    1052 GlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsn 1071
           ||||||||||||||||||||||||         ||||||   :::||||||      ||||||
Db    3727 GGAGACAACAGCCGTGTCGTGTCTCATAAGGAGATCGTGAAGGCAGCCAAGGAGGCCAAC 3786

Qy    1072 IleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGly 1091
           ||||||   ||||||:::   ||||||||||||||||||||:::||||||   ||||||
Db    3787 ATCCACCAGTTCATCGACTCACTGCCTGAGAAATACAACACCAGAGTGGGAGACAAAGGG 3846

Qy    1092 AlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLys 1111
              ||||||||||||||||||||||||:::||||||||||||||||||:::::::::
Db    3847 ACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATCGCCATCGCGCGCGCCCTCGTCAGACAG 3906

Qy    1112 ProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysVal 1131
           |||   ||||||||||||||||||||||||||||||||||      :::||||||||||||
Db    3907 CCTCACATCTTACTTCTGGATGAAGCGACATCAGCTCTGGATACGGAGAGTGAAAAGGTC 3966

Qy    1132 ValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArg 1151
           ||||||   ||||||||||||||||||   |||||||||||||:::|||:::   ||||||
Db    3967 GTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACCTGCATTGTGATCGCGCACCGC 4026

Qy    1152 LeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGlu 1171
           ||||||   ||||||||||||||||||||||||||||:::   ||||||::::::||||||
Db    4027 CTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATTCAGAACGGCCAGGTCAAGGAG 4086

Qy    1172 GlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 1191
              ||||||||||||:::||||||         :::   |||||||||   :::|||:::
Db    4087 CACGGCACCCACCAGCAGCTGCTGGCCCAGAAAGGCATCTATTTCTCGATGGTCAGTGTG 4146

Qy    1192 GlnSer 1193
           |||:::
Db    4147 CAGGCT 4152
```

RESULT 12
US-10-072-621-2
; Sequence 2, Application US/10072621
; Patent No. US20020169137A1
; GENERAL INFORMATION:
; APPLICANT: Reiner, Peter B.
; APPLICANT: Connop, Bruce P.
; APPLICANT: Pollard, Michelle
; TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION
; TITLE OF INVENTION: BY MODIFICATION OF ABC TRANSPORTER EXPRESSION OR ACTIVITY
; FILE REFERENCE: 100103.402
; CURRENT APPLICATION NUMBER: US/10/072,621
; CURRENT FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 4643
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-072-621-2

```
Alignment Scores:
Pred. No.:              3.21e-309      Length:        4643
Score:                  3537.50        Matches:       681
Percent Similarity:     75.55%         Conservative:  246
Best Local Similarity:  55.50%         Mismatches:    263
Query Match:            58.84%         Indels:        37
DB:                     9              Gaps:          10

US-09-873-409-6 (1-1195) x US-10-072-621-2 (1-4643)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||::::::|||   ||||||::::::::::::::||||||   ||||||||||||   |||
Db     575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy      21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
```

```
Db    635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db    695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    875 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    935 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAAATT 994

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db   1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db   1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db   1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db   1235 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1294

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db   1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1354

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db   1355 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1408

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db   1409 GGACAAGTACTCACTGTA---TTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1465

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db   1466 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1525

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db   1526 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1585

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db   1586 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1645

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   1646 ATCTTGAAGGGTCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1705

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   1706 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1765

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   1766 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1825

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   1826 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1885

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1886 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1945

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1946 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2005

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   2006 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2065

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   2066 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2125

Qy    501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db   2126 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 2185

Qy    521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db   2186 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 2245

Qy    541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db   2246 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 2305

Qy    561 Gln-----------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
Db   2306 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 2365

Qy    575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 589
Db   2366 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2425

Qy    590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
Db   2426 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2485

Qy    608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627
Db   2486 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2545

Qy    628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
Db   2546 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2605

Qy    648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
Db   2606 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2665

Qy    667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
Db   2666 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2725

Qy    687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
Db   2726 ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2785

Qy    707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
Db   2786 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2845
```

```
Qy       727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
             |||:::|||||| ||||||||| |||||| ||| |||||::::::||||
Db      2846 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2905

Qy       747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
                |||||||||::: |||:::||||||||| |||:::||| :::|||
Db      2906 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2965

Qy       767 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
             ||||||||||||||||||||:::::::||| |||:::|||:::||| |||::::::
Db      2966 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 3025

Qy       787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
             |||:::   |||::::::|||     ::::::||| ||| |||||||||:::
Db      3026 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 3085

Qy       807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
             ||||||:::  ||||||||||||||||||||||||:::.|||||| ||||||:::|||
Db      3086 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 3145

Qy       827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
             |||||||||:::|||:::  |||||| |||||| ::: ||||||     :::
Db      3146 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 3205

Qy       847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
             ||||||:::  :::|||||| ||| |||  ::::::|||:::  |||
Db      3206 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 3265

Qy       867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
             :::||||||:::||||||  ||||||||||||||||||||:::
Db      3266 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 3325

Qy       887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
             |||:::  |||  :::  :::||||||:::|||::: :::|||||||||:::
Db      3326 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 3385

Qy       907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
             |||:::          ||||||:::|||:::|||||||| |||||||||:::
Db      3386 GGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3445

Qy       927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
                 ::::::||||| ||| ||||||||| ||| ||||||   |||
Db      3446 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3505

Qy       947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
             :::|||  |||||||||::: ||| |||||| ||| |||||| ||||||
Db      3506 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3565

Qy       967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
             |||:::   :::|||:::||||||||||   :::::::::|||:::|||:::|||
Db      3566 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3625

Qy       987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
             ||||||||||||||||||||||||||||  ||||||||||||:::||| ||||||
Db      3626 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3685

Qy      1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
             |||:::  |||:::|||||| |||||| ::: ||| |||||||||||||||
Db      3686 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3745

Qy      1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
             ||||||:::  :::  |||||| |||||||||:::||||||:::||||||||||||
Db      3746 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3805

Qy      1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
             |||||||||||||||||||||||||||||||||||||  :::||||||   |||
Db      3806 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATTGTGAGGGCA 3865

Qy      1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
             |||   ||||||||||||:::||||||||| |||||| ||||||:::|||:::
Db      3866 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3925

Qy      1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
             |||||| |||||| ||||||||||||||||||||||||||:::||||||||||||
Db      3926 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3985

Qy      1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
             ||||||:::::::::::||| ||||||||||||||||||||||||||||||||||
Db      3986 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 4045

Qy      1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
             :::||||||||||||||||||| |||||||||||||||||| |||||||||||:::
Db      4046 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 4105

Qy      1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
             |||:::  |||||||||||| |||||||||||||||||||||||||||     |||
Db      4106 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 4165

Qy      1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
             |||::::::|||||| ||||||||||||:::||||||     :::  |||||||||
Db      4166 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 4225

Qy      1187 LysLeuValAsnAlaGlnSer 1193
                :::|||:::  |||:::
Db      4226 TCAATGGTCAGTGTCCAGGCT 4246
```

RESULT 13
US-09-866-866A-5
; Sequence 5, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 4189
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-866-866A-5

```
Alignment Scores:
Pred. No.:              2.71e-308     Length:        4189
Score:                  3526.50       Matches:       678
Percent Similarity:     74.74%        Conservative:  242
Best Local Similarity:  55.08%        Mismatches:    262
Query Match:            58.66%        Indels:        49
DB:                     10            Gaps:          10

US-09-873-409-6 (1-1195) x US-09-866-866A-5 (1-4189)

Qy         1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
             ||||||||||||  ||||||::::::::::::||| |||||||||:::  |||
Db       148 ATGATTCTGGGAACTCTCGCTGCTATTATCCATGGAACATTACTTCCCCTCTTGATGCTG 207

Qy        21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
             |||   |||  |||:::|||:::
Db       208 GTGTTTGGAAACATGACAGATAGTTTTACAAAAGCAGAAGCCAGTATTCTGCCAAGCATT 267

Qy        31 ------------------SerGlyCysLeuValGlnThrAsnThr--------------- 39
                               :::    |||:::  :::|||:::
Db       268 ACTAATCAAAGTGGACCCAACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAAGAG 327

Qy        40 ---------TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56
                      |||::::::               ||||||   |||||||||
```

```
Db    328 ATGGCCATATACGCCTAC---------------TATTACACCGGGATTGGTGCTGGTGTG 372
Qy     57 LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys 76
Db    373 CTCATAGTTGCCTACATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGAAGACAGATACAC 432
Qy     77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96
Db    433 AAGATTAGGCAGAAGTTTTTCCATGCTATAATGAATCAGGAGATAGGCTGGTTTGATGTG 492
Qy     97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
Db    493 CATGATGTTGGGGAGCTCAACACCCGGCTCACAGATGATGTCTCCAAAATTAATGACGGA 552
Qy    116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
Db    553 ATTGGTGACAAAATTGGGATGTTTTTCAGTCCATAACCACATTTTTAGCCGGTTTTATC 612
Qy    136 ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
Db    613 ATAGGATTTATAAGTGGTTGGAAGCTAACCCTTGTCATTTTGGCTGTCAGCCCTCTTATT 672
Qy    156 MetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
Db    673 GGATTGTCATCTGCTTTGTGGGCAAAGGTATTGACTTCATTTACTAATAAGGAACTCCAG 732
Qy    176 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIle 195
Db    733 GCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT 792
Qy    196 AlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLys 215
Db    793 GCCTTTGGAGGACAACAGAAGGAACTTGAAAGGTACAATAAAAATTTAGAAGAAGCTAAA 852
Qy    216 AspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePhe 235
Db    853 AATGTTGGCATAAAGAAAGCTATCACAGCCAGCATTTCGATAGGCATTGCCTACCTGTTG 912
Qy    236 MetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGlu 255
Db    913 GTCTATGCATCATATGCACTGGCATTCTGGTATGGGACATCCTTGGTCCTCTCAAATGAA 972
Qy    256 ProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyr 275
Db    973 ------TATTCTATTGGAGAAGTGCTTACTGTCTTCTTCTCTATTTTGTTGGGGACTTTT 1026
Qy    276 CysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPhe 295
Db   1027 AGTATTGGACACTTGGCCCCAAACATAGAAGCCTTTGCAAACGCACGAGGGGCAGCCTTT 1086
Qy    296 HisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyr 315
Db   1087 GAAATCTTCAAGATAATTGATAACGAGCCAAGCATTGACAGCTTCTCAACAAAGGGCTAC 1146
Qy    316 LysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSer 335
Db   1147 AAACCAGACAGTATAATGGGAAACTTAGAGTTTAAAAATGTTCACTTCAACTACCCATCG 1206
Qy    336 ArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrVal 355
Db   1207 AGAAGCGAAGTTCAGATCTTGAAGGGCCTCAATCTGAAGGTGAAGAGCGGACAGACGGTG 1266
Qy    356 AlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeu 375
Db   1267 GCCTTGGTTGGCAACAGTGGCTGTGGAAAAAGCACAACTGTCCAGCTGATGCAGAGGCTC 1326
Qy    376 TyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVal 395
Db   1327 TACGACCCCCTGGAGGGCGTGGTCAGTATCGACGGACAAGACATCAGAACCATCAATGTG 1386
Qy    396 ArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThr 415
Db   1387 AGGTATCTGAGGGAGATCATTGGTGTGGTGAGTCAGGAACCTGTGCTGTTTGCCACCACG 1446

Qy    416 IleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAla 435
Db   1447 ATCGCCGAGAACATTCGCTATGGCCGAGAAGATGTCACCATGGATGAGATTGAGAAAGCT 1506
Qy    436 AlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeu 455
Db   1507 GTCAAGGAAGCCAATGCCTATGACTTCATCATGAAACTGCCCCACCAATTTGACACCCTG 1566
Qy    456 ValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArg 475
Db   1567 GTTGGTGAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGAATCGCCATTGCCCGG 1626
Qy    476 AlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSer 495
Db   1627 GCCCTGGTCCGCAATCCCAAGATCCTTTTGTTGGACGAGGCCACCTCAGCCCTGGATACA 1686
Qy    496 GluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIle 515
Db   1687 GAAAGTGAAGCTGTGGTGCAGGCCGCACTGGATAAGGCTAGAGAAGGCCGGACCACCATT 1746
Qy    516 ValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAsp 535
Db   1747 GTGATAGCTCATCGCTTGTCTACAGTTCGTAATGCTGACGTCATTGCTGGTTTTGATGGT 1806
Qy    536 GlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyr 555
Db   1807 GGTGTCATTGTGGAGCAAGGAAATCATGATGAGCTCATGAGAGAAAAGGGCATTTACTTC 1866
Qy    556 SerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyr 575
Db   1867 AAACTTGTCATGACACAG---ACTAGAGGAAATGAAATTGAACCAGGAAATAATGCTTAT 1923
Qy    576 SerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAsp 595
Db   1924 GGATCCCAGAGTGACACTGATGCTTCTGAACTGACTTCAGAAGAATCC---AAATCACCT 1980
Qy    596 PheIleAspLysAla------------------GluGluSerThrGlnSerLysGluIle 609
Db   1981 TTAATAAGGAGATCAATTTACAGAAGTGTCCACAGAAAGCAAGACCAAGAGAGAAGACTT 2040
Qy    610 SerLeu------------------------ProGluValSerLeuLeuLysIleLeuLys 621
Db   2041 AGTATGAAAGAGGCTGTGGATGAAGATGTGCCTCTGGTTTCCTTTTGGCGGATCCTAAAT 2100
Qy    622 LeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGly 641
Db   2101 CTAAATCTAAGTGAATGGCCTTATTTACTTGTTGGCGTACTTTGCGCTGTTATAAATGGG 2160
Qy    642 ThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGly---Asn 660
Db   2161 TGCATACAACCAGTGTTTGCCATAGTATTTTCAAGGATTGTAGGGGTTTTTTCAAGAGAT 2220
Qy    661 AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeu 680
Db   2221 GATGACCATGAAACTAAACGACAGAATTGTAATTTGTTTTCCCTGTTCTTTCTGGTTATG 2280
Qy    681 GlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGlu 700
Db   2281 GGGCTGATTTCTTTTGTTACATATTTCTTTCAGGGCTTCACATTTGGCAAAGCCGGAGAG 2340
Qy    701 IleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAla 720
Db   2341 ATCCTCACCAAGCGAGTCCGATACATGGTTTTCAAATCCATGCTGAGACAGGATATAAGC 2400
Qy    721 TrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIle 740
Db   2401 TGGTTCGATGACCATAAGAACAGCACTGGCTCACTGACCACCAGGCTCGCCAGTGATGCT 2460
Qy    741 AlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsn 760
Db   2461 TCTAGTGTTAAAGGGGCGATGGGCGCCAGGCTTGCTGTAGTTACCCAGAATGTAGCAAAC 2520
```

```
Qy    761 MetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeu 780
Db   2521 CTCGGGACAGGAGTCATCCTCTCCTTAGTCTATGGCTGGCAGCTGACACTTCTACTTGTA 2580

Qy    781 SerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPhe 800
Db   2581 GTAATTATACCGCTCATTGTATTGGGCGGAATTATTGAAATGAAGCTGTTGTCTGGCCAA 2640

Qy    801 AlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGlu 820
Db   2641 GCCTTGAAGGACAAGAAACAGCTTGAGATCTCTGGGAAGATTGCTACAGAAGCAATTGAA 2700

Qy    821 AsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGlu 840
Db   2701 AACTTCCGCACTATTGTCTCTTTGACTCGGGAGCAGAAGTTTGAAACCATGTATGCCCAG 2760

Qy    841 MetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyr 860
Db   2761 AGCTTGCAGGTACCATACAGAAATGCGATGAAGAAAGCACACGTGTTTGGGATCACGTTC 2820

Qy    861 AlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyr 880
Db   2821 TCCTTCACCCAGGCCATGATGTATTTTTCTTATGCTGCTTGTTTCCGGTTCGGTGCCTAC 2880

Qy    881 LeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAla 900
Db   2881 TTGGTGGCACAACAACTCATGACTTTTGAAAATGTTATGTTGGTATTTTCTGCTGTTGTC 2940

Qy    901 TyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLys 920
Db   2941 TTTGGTGCCATGGCAGCTGGGAATACTAGTTCATTTGCTCCTGACTATGCGAAAGCCAAA 3000

Qy    921 SerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSer 940
Db   3001 GTATCAGCATCTCATATCATCAGGATCATTGAGAAAACCCCTGAGATTGACAGCTACAGC 3060

Qy    941 GlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhe 960
Db   3061 ACAGAGGGCTTGAAGCCTACTCTGTTAGAAGGAAATGTAAAATTTAATGGAGTCCAGTTT 3120

Qy    961 PheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArg 980
Db   3121 AACTATCCCACCCGACCCAACATCCCAGTGCTTCAGGGGCTGAGCCTCGAGGTGAAGAAG 3180

Qy    981 GlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeu 1000
Db   3181 GGCCAGACGTTGGCCCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTG 3240

Qy   1001 LeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLys 1020
Db   3241 CTCGAGCGCTTCTACGACCCCATGGCTGGATCAGTGTTTCTAGATGGCAAAGAAATAAAG 3300

Qy   1021 GluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeu 1040
Db   3301 CAACTGAATGTCCAGTGGCTCCGAGCTCACCTTGGCATTGTGTCCCAGGAGCCCATTCTC 3360

Qy   1041 PheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeu 1060
Db   3361 TTTGACTGCAGCATTGCAGAGAACATCGCCTATGGAGACAACAGCCGGGCCGTGTCTCAT 3420

Qy   1061 AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuPro 1080
Db   3421 GAGGAGATTGTGAGGGCAGCCAAGGAGGCCAACATCCACCAGTTCATCGACTCACTGCCT 3480

Qy   1081 GluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGln 1100
Db   3481 GATAAATACAACACCAGAGTAGGAGACAAAGGCACTCAGCTGTCGGGTGGGCAGAAGCAG 3540

Qy   1101 ArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAla 1120
Db   3541 CGCATCGCCATCGCACGTGCCCTCGTCAGACAGCCTCACATTTTACTTCTGGACGAAGCA 3600

Qy   1121 ThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArg 1140
Db   3601 ACATCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAGGAAGCGCTGGACAAAGCCAGG 3660

Qy   1141 ThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeu 1160
Db   3661 GAAGGCCGCACCTGCATTGTGATCGCTCACCGCCTGTCCACCATCCAGAACGCGGACTTG 3720

Qy   1161 IleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArg 1180
Db   3721 ATCGTGGTGATTGAGAACGGCAAAGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCG 3780

Qy   1181 AsnArgAspIleTyrPheLysLeuValAsnAla 1191
Db   3781 CAGAAGGGCATCTACTTCTCAATGGTCCAGGCT 3813
```

RESULT 14
US-09-866-866A-7
; Sequence 7, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
;   LENGTH: 4788
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-866-866A-7

Alignment Scores:			
Pred. No.:	7.58e-308	Length:	4788
Score:	3522.50	Matches:	668
Percent Similarity:	74.74%	Conservative:	252
Best Local Similarity:	54.26%	Mismatches:	264
Query Match:	58.59%	Indels:	47
DB:	10	Gaps:	8

US-09-873-409-6 (1-1195) x US-09-866-866A-7 (1-4788)

```
Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    148 ATGCTGGTGGGAACTCTGGCTGCTATTATCCATGGAGTGGCGCTCCCACTTATGATGCTG 207

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsn------ 38
Db    208 ATCTTTGGTGACATGACAGATAGCTTTGCAAGTGTAGGAAACGTCTCTAAAAACAGTACT 267

Qy     39 ---------------------------------------------------------Thr 39
Db    268 AATATGAGTGAGGCCGATAAAAGAGCCATGTTTGCCAAACTGGAGGAAGAAATGACCACG 327

Qy     40 TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db    328 TACGCCTAC---------------TATTACACCGGGATTGGTGCTGGTGTGCTCATAGTT 372

Qy     60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db    373 GCCTACATCCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAGATACACAAGATCAGG 432

Qy     80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db    433 CAGAAGTTTTTTCATGCTATAATGAATCAGGAGATAGGCTGGTTTGATGTGCATGACGTT 492
```

```
Qy    100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db    493 GGGGAGCTCAACACCCGGCTCACAGATGATGTTTCCAAAATTAATGAAGGAATTGGTGAC 552

Qy    119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db    553 AAAATCGGAATGTTCTTCCAGGCAATGGCAACATTTTTGGTGGTTTATAATAGGATTT 612

Qy    139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db    613 ACCCGTGGCTGGAAGCTAACCCTTGTGATTTGGCCATCAGCCCTGTTCTTGGACTGTCA 672

Qy    159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db    673 GCTGGTATTTGGGCAAAGATATTGTCTTCATTTACTGATAAGGAACTCCATGCTTATGCA 732

Qy    179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db    733 AAAGCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCGTTTGGA 792

Qy    199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
Db    793 GGACAAAAGAAGGAACTTGAAAGGTACAATAACAACTTGGAAGAAGCTAAAAGGCTGGGG 852

Qy    219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238
Db    853 ATAAAGAAAGCTATCACGGCCAACATCTCCATGGGTGCAGCTTTTCTCCTTATCTATGCA 912

Qy    239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
Db    913 TCATATGCTCTGGCATTCTGGTATGGGACTTCCTTGGTCATCTCCAAAGAA------TAC 966

Qy    259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
Db    967 TCTATTGGACAAGTGCTCACTGTCTTCTTTTCCGTGTTAATTGGAGCATTCAGTGTTGGA 1026

Qy    279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
Db   1027 CAGGCATCTCCAAATATTGAAGCCTTCGCCAATGCACGAGGAGCAGCTTATGAAGTCTTC 1086

Qy    299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
Db   1087 AAAATAATTGATAATAAGCCCAGTATAGACAGCTTCTCAAAGAGTGGGCACAAACCAGAC 1146

Qy    319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
Db   1147 AACATACAAGGAAATCTGGAATTTAAGAATATTCACTTCAGTTACCCATCTCGAAAAGAA 1206

Qy    339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
Db   1207 GTTCAGATCTTGAAGGGCCTCAATCTGAAGGTGAAGAGCGGACAGACGGTGGCCCTGGTT 1266

Qy    359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
Db   1267 GGCAACAGTGGCTGTGGAAAAAGCACAACTGTCCAGCTGATGCAAAGGCTCTACGACCCC 1326

Qy    379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
Db   1327 CTAGATGGCATGGTCAGTATCGACGGACAGGACATCAGAACCATCAATGTGAGGTATCTG 1386

Qy    399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
Db   1387 AGGGAGATCATTGGTGTGGTGAGTCAGGAACCTGTGCTGTTTGCCACCACGATCGCCGAG 1446

Qy    419 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
Db   1447 AACATTCGCTATGGCCGAGAAGATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAA 1506

Qy    439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
Db   1507 GCCAATGCCTATGACTTCATCATGAAACTGCCCCACCAATTTGACACCCTGGTTGGTGAG 1566

Qy    459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
Db   1567 AGAGGGGCGCACGTGAGTGGGGGACAGAAACAGAGAATCGCCATTGCCCGGGCCCTGGTC 1626

Qy    479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
Db   1627 CGCAATCCCAAGATCCTTTTGTTGGACGAGGCCACCTCAGCCCTGGATACAGAAAGTGAA 1686

Qy    499 SerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 518
Db   1687 GCTGTGGTTCAGGCCGCACTGGATAAGGCTAGAGAAGGCCGGACCACCATTGTGATAGCT 1746

Qy    519 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 538
Db   1747 CATCGCTTGTCTACCGTTCGTAATGCTGACGTCATTGCTGGTTTTGATGGTGGTGTCATT 1806

Qy    539 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 558
Db   1807 GTGGAGCAAGGAAATCATGATGAGCTCATGAGAGAAAAGGGCATTTACTTCAAACTTGTC 1866

Qy    559 MetSerGln-----------------------------AspIleLysLysAlaAspGlu 568
Db   1867 ATGACACAGACAGCAGGAAATGAAATTGAATTAGGAAATGAAGCTTGTAAATCTAAGGAT 1926

Qy    569 GlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSer 588
Db   1927 GAAATTGATAATTTAGACATGTCTTCAAAAGATTCAGGATCCAGTCTAATAAGAAGA--- 1983

Qy    589 ValLysSerIleLysSerAspPheIle---------------AspLysAlaGluGluSer 603
Db   1984 ---AGATCAACTCGCAAAAGCATCTGTGGACCACATGACCAAGACAGGAAGCTTAGTACC 2040

Qy    604 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 623
Db   2041 AAAGAGGCCCTGGATGAAGATGTACCTCCAGCTTCCTTTTGGCGGATCCTGAAGTTGAAT 2100

Qy    624 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 643
Db   2101 TCAACTGAATGGCCTTATTTTGTGGTTGGTATATTCTGTGCCATAATAAATGGAGGCTTA 2160

Qy    644 HisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAsp 662
Db   2161 CAGCCAGCATTCTCCGTAATATTTTCAAAAGTTGTAGGGGTTTTTACAAATGGTGGCCCC 2220

Qy    663 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 682
Db   2221 CCTGAAACCCAGCGGCAGAACAGCAACTTGTTTTCCTTGTTGTTTCTGATCCTTGGGATC 2280

Qy    683 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 702
Db   2281 ATTTCTTTCATTACATTTTTTCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTC 2340

Qy    703 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 722
Db   2341 ACCAAGCGACTCCGATACATGGTTTTCAAATCCATGCTGAGACAGGATGTGAGCTGGTTT 2400

Qy    723 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 742
Db   2401 GATGACCCTAAAAACACCACCGGAGCACTGACCACCAGGCTCGCCAACGATGCTGCTCAA 2460

Qy    743 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 762
Db   2461 GTGAAAGGGGCTACAGGGTCTAGGCTTGCTGTGATTTTCCAGAACATAGCAAATCTTGGG 2520

Qy    763 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 782
Db   2521 ACAGGAATCATCATATCCCTAATCTATGGCTGGCAACTAACACTTTTACTCTTAGCAATT 2580

Qy    783 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 802
Db   2581 GTACCCATCATTGCGATAGCTGGAGTGGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 2640

Qy    803 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 822
```

```
Db   2641 AAAGATAAGAAGGAACTAGAAGGTTCTGGAAAGATTGCTACGGAAGCAATTGAAAACTTC 2700

Qy    823 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 842

Db   2701 CGCACTGTTGTCTCTTTGACTCGGGAGCAGAAGTTTGAAACCATGTATGCCCAGAGCTTG 2760

Qy    843 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 862

Db   2761 CAGATACCATACAGAAATGCGATGAAGAAAGCACACGTGTTTGGGATCACGTTCTCCTTC 2820

Qy    863 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 882

Db   2821 ACCCAGGCCATGATGTATTTTTCTTATGCTGCTTGTTTCCGGTTCGGTGCCTACTTGGTG 2880

Qy    883 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 902

Db   2881 ACACAACAACTCATGACTTTTGAAAATGTTCTGTTAGTATTCTCAGCTATTGTCTTTGGT 2940

Qy    903 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 922

Db   2941 GCCATGGCAGTGGGGCAGGTCAGTTCATTCGCTCCTGACTATGCGAAAGCAACAGTGTCA 3000

Qy    923 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 942

Db   3001 GCATCCCACATCATCAGGATCATTGAGAAAACCCCCGAGATTGACAGCTACAGCACGCAA 3060

Qy    943 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 962

Db   3061 GGCCTAAAGCCGAATATGTTGGAAGGAAATGTGCAATTTAGTGGAGTCGTGTTCAACTAT 3120

Qy    963 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 982

Db   3121 CCCACCCGACCCAGCATCCCAGTGCTTCAGGGGCTGAGCCTTGAGGTGAAGAAGGGCCAG 3180

Qy    983 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 1002

Db   3181 ACGCTGGCCCTGGTGGGCAGCAGTGGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAG 3240

Qy   1003 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 1022

Db   3241 CGCTTCTACGACCCCATGGCTGGATCAGTGTTTCTAGATGGCAAAGAAATAAAGCAACTG 3300

Qy   1023 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 1042

Db   3301 AATGTCCAGTGGCTCCGAGCACAGCTGGGCATTGTGTCCCAAGAGCCCATTCTCTTTGAC 3360

Qy   1043 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 1062

Db   3361 TGCAGCATCGCAGAGAACATTGCCTACGGAGACAACAGCCGGGTCGTGTCTTATGAGGAG 3420

Qy   1063 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 1082

Db   3421 ATTGTGAGGGCAGCCAAGGAGGCCAACATCCACCAGTTCATCGACTCGCTACCTGATAAA 3480

Qy   1083 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1102

Db   3481 TACAACACCAGAGTAGGAGACAAAGGCACTCAGCTGTCGGGTGGGCAGAAGCAGCGCATC 3540

Qy   1103 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 1122

Db   3541 GCCATCGCACGCGCCCTCGTCAGACAGCCTCACATTTTACTTCTGGACGAAGCAACATCA 3600

Qy   1123 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1142

Db   3601 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGC 3660

Qy   1143 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1162

Db   3661 CGCACCTGCATTGTGATCGCTCACCGCCTGTCCACCATCCAGAACGCGGACTTGATCGTG 3720

Qy   1163 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1182

Db   3721 GTGATTCAGAACGGCAAGGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCGCAGAAG 3780

Qy   1183 AspIleTyrPheLysLeuValAsnAlaGlnSer 1193

Db   3781 GGCATCTACTTCTCAATGGTCAGTGTGCAGGCT 3813
```

RESULT 15
US-09-880-107-2299
; Sequence 2299, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2299
;   LENGTH: 3924
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020142981A1 M23234
US-09-880-107-2299

Alignment Scores:

Pred. No.:	2.38e-306	Length:	3924
Score:	3504.50	Matches:	676
Percent Similarity:	74.43%	Conservative:	235
Best Local Similarity:	55.23%	Mismatches:	276
Query Match:	58.29%	Indels:	37
DB:	10	Gaps:	8

US-09-873-409-6 (1-1195) x US-09-880-107-2299 (1-3924)

```
Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20

Db    201 ATGTCGCTGGGTACCATCATGGCCATAGCTCACGGATCAGGTCTCCCCCTCATGATGATA 260

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40

Db    261 GTATTTGGAGAGATGACTGACAAATTTGTTGAT-----------ACTGCAGGAAACTTC 308

Qy     41 SerPhe---------PheArgLeuThrLeu----------------------------- 47

Db    309 TCCTTTCCAGTGAACTTTTCCTTGTCGCTGCTAAATCCAGGCAAAATTCTGGAAGAAGAA 368

Qy     48 -----------------TyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61

Db    369 ATGACTAGATATGCATATTACTACTCAGGATTGGGTGCTGGAGTTCTTGTTGCTGCCTAT 428

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81

Db    429 ATACAAGTTTCATTTTGGACTTTGGCAGCTGGTCGACAGATCAGGAAAATTAGGCAGAAG 488

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101

Db    489 TTTTTTCATGCTATTCTACGACAGGAAATAGGATGGTTTGACATCAATGACACCACTGAA 548

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120

Db    549 CTCAATACGCGGCTAACAGATGACATCTCCAAAATCAGTGAAGGAATTGGTGACAAGGTT 608

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
```

```
Db    609 GGAATGTTCTTTCAAGCAGTAGCCACGTTTTTGCAGGATTCATAGTGGGATTCATCAGA 668

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    669 GGATGGAAGCTCACCCTTGTGATAATGGCCATCAGCCCTATTCTAGGACTCTCTGCAGCC 728

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    729 GTTTGGGCAAAGATACTCTCGGCATTTAGTGACAAAGAACTAGCTGCTTATGCAAAAGCA 788

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    789 GGCGCCGTGGCAGAAGAGGCTCTGGGGGCCATCAGGACTGTGATAGCTTTCGGGGGCCAG 848

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db    849 AACAAAGAGCTGGAAAGGTATCAGAAACATTTAGAAAATGCCAAAGAGATTGGAATTAAA 908

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db    909 AAAGCTATTTCAGCAAACATTTCCATGGGTATTGCCTTCCTGTTAATATATGCATCATAT 968

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db    969 GCACTGGCCTTCTGGTATGGATCCACTCTAGTCATATCAAAAGAA------TATACTATT 1022

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db   1023 GGAAATGCAATGACAGTTTTTTTTTCAATCCTAATTGGAGCTTTCAGTGTTGGCCAGGCT 1082

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db   1083 GCCCCATGTATTGATGCTTTTGCCAATGCAAGAGGAGCAGCATATGTGATCTTTGATATT 1142

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db   1143 ATTGATAATAATCCTAAAATTGACAGTTTTTCAGAGAGAGGACACAAACCAGACAGCATC 1202

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db   1203 AAAGGGAATTTGGAGTTCAATGATGTTCACTTTTCTTACCCTTCTCGAGCTAACGTCAAG 1262

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   1263 ATCTTGAAGGGCCTCAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAGT 1322

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   1323 AGTGGCTGTGGGAAGAGCACAACGGTCCAGCTGATACAGAGGCTCTATGACCCTGATGAG 1382

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   1383 GGCACAATTAACATTGATGGGCAGGATATTAGGAACTTTAATGTAAACTATCTGAGGGAA 1442

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   1443 ATCATTGGTGTGGTGAGTCAGGAGCCGGTGCTGTTTTCCACCACAATTGCTGAAAATATT 1502

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1503 TGTTATGGCCGTGGAAATGTAACCATGGATGAGATAAAGAAAGCTGTCAAAGAGGCCAAC 1562

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1563 GCCTATGAGTTTATCATGAAATTACCACAGAAATTTGACACCCTGGTTGGAGAGAGAGGG 1622

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   1623 GCCCAGCTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1682

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   1683 CCCAAGATCCTTCTGCTGGATGAGGCCACGTCAGCATTGGACACAGAAAGTGAAGCTGAG 1742

Qy    501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db   1743 GTACAGGCAGCTCTGGATAAGGCCAGAGAAGGCCGGACCACCATTGTGATAGCACACCGA 1802

Qy    521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db   1803 CTGTCTACGGTCCGAAATGCAGATGTCATCGCTGGGTTTGAGGATGGAGTAATTGTGGAG 1862

Qy    541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal------ 558
Db   1863 CAAGGAAGCCACAGCGAACTGATGAAGAAGGAAGGGGTGTACTTCAAACTTGTCAACATG 1922

Qy    559 --------------------MetSerGlnAspIleLysLysAlaAspGluGlnMetGlu 571
Db   1923 CAGACATCAGGAAGCCAGATCCAGTCAGAAGAATTTGAACTAAATGATGAAAAGGCTGCC 1982

Qy    572 SerMetThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSer 591
Db   1983 ACTAGAATGGCCCCAAATGGCTGGAAATCTCGCCTATTTAGGCATTCTACTCAGAAAAAC 2042

Qy    592 IleLys---SerAspPheIleAspLysAla---GluGluSerThrGlnSerLysGluIle 609
Db   2043 CTTAAAAATTCACAAATGTGTCAGAAGAGCCTTGATGTGGAAACCGATGGACTTGAAGCA 2102

Qy    610 SerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPhe 629
Db   2103 AATGTGCCACCAGTGTCCTTTCTGAAGGTCCTGAAACTGAATAAAACAGAATGGCCCTAC 2162

Qy    630 ValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIle 649
Db   2163 TTTGTCGTGGGAACAGTATGTGCCATTGCCAATGGGGGGCTTCAGCCGGCATTTTCAGTC 2222

Qy    650 IlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAsp 669
Db   2223 ATATTCTCAGAGATCATAGCGATTTTTGGACCAGGCGATGATGCAGTGAAGCAGCAGAAG 2282

Qy    670 AlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPhe 689
Db   2283 TGCAACATATTCTCTTTGATTTTCTTATTTCTGGGAATTATTTCTTTTTTTACTTTCTTC 2342

Qy    690 MetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeu 709
Db   2343 CTTCAGGGTTTCACGTTTGGGAAAGCTGGCGAGATCCTCACCAGAAGACTGCGGTCAATG 2402

Qy    710 AlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThr 729
Db   2403 GCTTTTAAAGCAATGCTAAGACAGGACATGAGCTGGTTTGATGACCATAAAAACAGTACT 2462

Qy    730 GlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySer 749
Db   2463 GGTGCACTTTCTACAAGACTTGCCACAGATGCTGCCCAAGTCCAAGGAGCCACAGGAACC 2522

Qy    750 ArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPhe 769
Db   2523 AGGTTGGCTTTAATTGCACAGAATATAGCTAACCTTGGAACTGGTATTATCATATCATTT 2582

Qy    770 IleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThr 789
Db   2583 ATCTACGGTTGGCAGTTAACCCTATTGCTATTAGCAGTTGTTCCAATTATTGCTGTGTCA 2642

Qy    790 GlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLys 809
Db   2643 GGAATTGTTGAAATGAAATTGTTGGCTGGAAATGCCAAAAGAGATAAAAAAGAACTGGAA 2702

Qy    810 HisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThr 829
Db   2703 GCTGCTGGAAAGATTGCAACAGAGGCAATAGAAAATATTAGGACAGTTGTGTCTTTGACC 2762

Qy    830 ArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThr 849
Db   2763 CAGGAAAGAAAATTTGAATCAATGTATGTTGAAAAATTGTATGGACCTTACAGGAATTCT 2822
```

```
Qy    850 SerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPhe 869
             :::||||||   |||   |||         ::::::   |||   ||||||:::||||||
Db   2823 GTGCAGAAGGCACACATCTATGGAATTACTTTTAGTATCTCACAAGCATTATGTATTTT 2882

Qy    870 AlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrPro 889
          :::||||||   ||||||||||||||||||||||||||||||||   |||   |||
Db   2883 TCCTATGCCGGTTGTTTTCGATTTGGTGCATATCTCATTGTGAATGGACATATGCGCTTC 2942

Qy    890 GluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThr 909
                :::   :::||||||:::||||||   :::||||||:::|||:::|||
Db   2943 AGAGATGTTATTCTGGTGTTTTCTGCAATTGTATTTGGTGCAGTGGCTCTAGGACATGCC 3002

Qy    910 LeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeu 929
                            ||||||:::|||:::|||||||||   |||||||||||||||   |||
Db   3003 AGTTCATTTGCTCCAGACTATGCTAAAGCTAAGCTGTCTGCAGCCCACTTATTCATGCTG 3062

Qy    930 LeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCys 949
             |||::::::|||   |||||||||   |||:::||||||   |||||||||
Db   3063 TTTGAAAGACAACCTCTGATTGACAGCTACAGTGAAGAGGGGCTGAAGCCTGATAAATTT 3122

Qy    950 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 969
          |||||||||:::   |||   ||||||   |||   ||||||   |||   :::|||
Db   3123 GAAGGAAATATAACATTTAATGAAGTCGTGTTCAACTATCCCACCCGAGCAAACGTGCCA 3182

Qy    970 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 989
          :::|||:::||||||||||||   :::::::::|||:::|||:::|||   |||||||||
Db   3183 GTGCTTCAGGGGCTGAGCCTGGAGGTGAAGAAAGGCCAGACACTAGCCCTGGTGGGCAGC 3242

Qy    990 SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGln 1009
          ||||||||||||||||||   ||||||||||||:::|||   |||||||||:::
Db   3243 AGTGGCTGTGGGAAGAGCACGGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCG 3302

Qy   1010 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 1029
          |||   ||||||   ||||||   :::||||||:::||||||||||||||||||||:::
Db   3303 GGGACAGTGCTTCTCGATGGTCAAGAAGCAAAGAAACTCAATGTCCAGTGGCTCAGAGCT 3362

Qy   1030 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 1049
          |||:::   ||||||   |||||||||:::||||||:::||||||||||||||||||||||
Db   3363 CAACTCGGAATCGTGTCTCAGGAGCCTATCCTATTTGACTGCAGCATTGCCGAGAATATT 3422

Qy   1050 AlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla 1069
          ||||||||||||||||||||||||||||   |||||||||   ||||||   |||
Db   3423 GCCTATGGAGACAACAGCCGGGTTGTATCACAGGATGAAATTGTGAGTGCAGCCAAAGCT 3482

Qy   1070 AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeu 1089
          ||||||||||||   |||||||||   ||||||   ||||||   |||:::||||||
Db   3483 GCCAACATACATCCTTTCATCGAGACGTTACCCCACAAATATGAAACAAGAGTGGGAGAT 3542

Qy   1090 LysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeu 1109
          ||||||   ||||||||||||||||||||||||||:::||||||||||||||||||:::
Db   3543 AAGGGGACTCAGCTCTCAGGAGGTCAAAAACAGAGGATTGCTATTGCCCGAGCCCTCATC 3602

Qy   1110 GlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGlu 1129
          ::::::|||:::|||||||||||||||||||||||||||||||||||||   :::||||||
Db   3603 AGACAACCTCAAATCCTCCTGTTGGATGAAGCTACATCAGCTCTGGATACTGAAAGTGAA 3662

Qy   1130 LysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThr 1149
          ||||||||||||   ||||||||||||||||||   ||||||||||||:::|||:::
Db   3663 AAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCT 3722

Qy   1150 HisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIle 1169
          ||||||||||||   ||||||||||||||||||||||||||||||   ||||||::::::
Db   3723 CACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAATGGGAGAGTC 3782

Qy   1170 LysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuVal 1189
          ||||||   ||||||||||||:::||||||      :::   ||||||||||||   :::|||
Db   3783 AAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTC 3842

Qy   1190 AsnAlaGlnSer 1193
          :::   |||:::
Db   3843 AGTGTCCAGGCT 3854
```

Search completed: March 31, 2003, 16:13:21
Job time : 686.971 secs

```
                GenCore version 5.1.4_p5_4578
          Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 30, 2003, 01:24:25 ; Search time 16155.1 Seconds
                                          (without alignments)
                                          2152.747 Million cell updates/sec

Title:          US-09-873-409-6
Perfect score:  6012
Sequence:       1 MILGILASLVNGACLPLMPL..........QELLRNRDIYFKLVNAQSVQ 1195

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters:      4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115418_19228/app_query.fasta_1.7544
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_54887_@runat_27032003_115418_19228 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      GenEmbl:*
                1:  gb_ba:*
                2:  gb_htg:*
                3:  gb_in:*
                4:  gb_om:*
                5:  gb_ov:*
                6:  gb_pat:*
                7:  gb_ph:*
                8:  gb_pl:*
                9:  gb_pr:*
                10:  gb_ro:*
                11:  gb_sts:*
                12:  gb_sy:*
                13:  gb_un:*
                14:  gb_vi:*
                15:  em_ba:*
                16:  em_fun:*
                17:  em_hum:*
                18:  em_in:*
                19:  em_mu:*
                20:  em_om:*
                21:  em_or:*
                22:  em_ov:*
                23:  em_pat:*
                24:  em_ph:*
                25:  em_pl:*
                26:  em_ro:*
                27:  em_sts:*
                28:  em_un:*
                29:  em_vi:*
                30:  em_htg_hum:*
                31:  em_htg_inv:*
                32:  em_htg_other:*
                33:  em_htg_mus:*
                34:  em_htg_pln:*
                35:  em_htg_rod:*
                36:  em_htg_mam:*
                37:  em_htg_vrt:*
                38:  em_sy:*
                39:  em_htgo_hum:*
                40:  em_htgo_mus:*
                41:  em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	6012	100.0	3621	6	AX339032	AX339032 Sequence
2	5988.5	99.6	3702	6	AX339031	AX339031 Sequence
3	5760.5	95.8	3699	6	AX478104	AX478104 Sequence
4	5159.5	85.8	3177	6	AX339030	AX339030 Sequence
5	4561.5	75.9	2856	6	AX339028	AX339028 Sequence
6	3574	59.4	4296	10	CRUPGPI	M60040 C.griseus P
7	3573	59.4	4304	10	CRUPGP1165	M59253 Chinese ham
8	3557	59.2	4279	6	AX105082	AX105082 Sequence
9	3556.5	59.2	4195	6	AX108656	AX108656 Sequence
10	3555.5	59.1	4927	10	AF257746	AF257746 Rattus no
11	3554	59.1	3860	6	AX322787	AX322787 Sequence
12	3554	59.1	4279	6	AX105078	AX105078 Sequence
13	3554	59.1	4279	6	AX105080	AX105080 Sequence
14	3554	59.1	4646	6	AX391099	AX391099 Sequence
15	3552	59.1	4045	12	AF269224	AF269224 Synthetic
16	3551	59.1	4378	6	E02326	E02326 Multidrug r
17	3549.5	59.0	3858	4	OAU78609	U78609 Ovis aries
18	3549	59.0	4279	6	AX105057	AX105057 Sequence
19	3549	59.0	4317	4	AF045016	AF045016 Canis fam
20	3549	59.0	4317	6	AX105059	AX105059 Sequence
21	3548	59.0	4186	6	AX108654	AX108654 Sequence
22	3545	59.0	3860	6	AX322789	AX322789 Sequence
23	3545	59.0	4646	6	AX336420	AX336420 Sequence
24	3545	59.0	4646	6	AX336708	AX336708 Sequence
25	3545	59.0	4646	6	I49610	I49610 Sequence 2
26	3545	59.0	4646	9	HUMMDR1	M14758 Homo sapien
27	3545	59.0	4669	6	I08557	I08557 Sequence 3
28	3545	59.0	6505	6	AR028671	AR028671 Sequence
29	3545	59.0	8630	6	AX012320	AX012320 Sequence
30	3545	59.0	8630	6	AX012321	AX012321 Sequence
31	3545	59.0	9318	6	AR028672	AR028672 Sequence
32	3543.5	58.9	4264	6	AR051647	AR051647 Sequence
33	3543.5	58.9	4264	6	AR051650	AR051650 Sequence
34	3541.5	58.9	4323	10	AF286167	AF286167 Rattus no
35	3537.5	58.8	3840	6	AX481416	AX481416 Sequence
36	3537.5	58.8	3988	6	AX024454	AX024454 Sequence
37	3537.5	58.8	4192	9	AF016535	AF016535 Homo sapi
38	3529	58.7	4280	10	CRUPGPII	M60041 C.griseus P
39	3527	58.7	4669	6	AR091275	AR091275 Sequence
40	3527	58.7	4669	6	AR203322	AR203322 Sequence
41	3526.5	58.7	4189	6	AX322791	AX322791 Sequence
42	3526.5	58.7	4298	10	MUSMDR	M14757 Mouse multi
43	3523	58.6	3934	4	CFA419568	AJ419568 Canis fam
44	3522.5	58.6	3888	10	AY082609	AY082609 Rattus no
45	3522.5	58.6	4788	6	AX322793	AX322793 Sequence

ALIGNMENTS

RESULT 1

```
AX339032
LOCUS       AX339032                3621 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 14 from Patent WO0194400.
ACCESSION   AX339032
VERSION     AX339032.1  GI:18129124
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 14 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..3621
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1081 a    696 c    842 g   1002 t
ORIGIN

Alignment Scores:
Pred. No.:              0              Length:         3621
Score:                  6012.00        Matches:        1195
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            100.00%        Indels:         0
DB:                     6              Gaps:           0

US-09-873-409-6 (1-1195) x AX339032 (1-3621)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       34 ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 93

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153

Qy       41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      154 TCTTTCTTCAGGTTGACCCTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGT 213

Qy       61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      214 TACATACAGATTTCCTTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAA 273

Qy       81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      274 CAGTTTTTTCATTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333

Qy      101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      334 GAACTTAACACTCGCATGACAGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATT 393

Qy      121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAG 453

Qy      141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      454 GGCTGGAAACTCACCCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA 513

Qy      161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573

Qy      181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAG 633

Qy      201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      634 GAGAAAGAACTTCAAAGGTATACACAGAATCTCAAAGATGCAAAGGATTTTGGCATAAAA 693

Qy      221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      694 AGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAATGGAACCTAT 753

Qy      241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      754 GGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGATATACCATC 813

Qy      261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      814 GGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATTGGAGCAGCA 873

Qy      281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      874 GTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTT 933

Qy      301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      934 ATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATA 993

Qy      321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      994 GAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAG 1053

Qy      341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1054 ATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTC 1113

Qy      361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1114 AATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGAT 1173

Qy      381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1174 GGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGAC 1233

Qy      401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1234 CATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATC 1293

Qy      421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1294 AAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAAT 1353

Qy      441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1354 GCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGA 1413

Qy      461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1414 GCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAAC 1473

Qy      481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1474 CCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCT 1533

Qy      501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1534 GTTCAAGCTGCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTAGCACACCGA 1593

Qy      521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1594 CTTTCTACTATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATGCTGGCGGAG 1653

Qy      541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1654 AAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTTGTGATGTCA 1713

Qy      561 GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 580
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1714 CAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAG 1773

Qy      581 ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAla 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1774 ACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCT 1833

Qy      601 GluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeu 620
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1834 GAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTA 1893

Qy      621 LysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 640
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1894 AAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCTGTTCTAAAT 1953

Qy      641 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsn 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1954 GGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATGTTTGGAAAT 2013

Qy      661 AsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeu 680
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2014 AATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTCGTCATTTTG 2073

Qy      681 GlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGlu 700
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2074 GGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGAGCAGGGGAA 2133

Qy      701 IleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAla 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2134 ATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAGGATATTGCC 2193

Qy      721 TrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIle 740
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2194 TGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCCATAGATATA 2253

Qy      741 AlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsn 760
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2254 GCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAATGCAACTAAC 2313

Qy      761 MetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeu 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2314 ATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTCCTGATTCTG 2373

Qy      781 SerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPhe 800
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2374 AGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATGACTGGATTT 2433

Qy      801 AlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGlu 820
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2434 GCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAG 2493

Qy      821 AsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGlu 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2494 AATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATGTATGAAGAG 2553

Qy      841 MetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyr 860
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2554 ATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTAT 2613

Qy      861 AlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyr 880
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2614 GCATTCAGCCATGCCTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTTGGAGCCTAT 2673

Qy      881 LeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAla 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2674 TTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACTGCAATTGCA 2733

Qy      901 TyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLys 920
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2734 TATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCCAAAGCCAAA 2793

Qy      921 SerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSer 940
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2794 TCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGACAGCCGCAGT 2853

Qy      941 GlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhe 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2854 CAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTC 2913

Qy      961 PheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArg 980
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2914 TTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGA 2973

Qy      981 GlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeu 1000
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     2974 GGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGAAAAAGCACTTCTGTTCAACTT 3033

Qy     1001 LeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLys 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3034 CTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTGGATGCAAAA 3093

Qy     1021 GluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeu 1040
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3094 GAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAGCCTGTGCTC 3153

Qy     1041 PheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeu 1060
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3154 TTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTGGTGCCATTA 3213

Qy     1061 AspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuPro 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3214 GATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAAGGTCTCCCT 3273

Qy     1081 GluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGln 1100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3274 GAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAACAA 3333

Qy     1101 ArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAla 1120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3334 AGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTGGATGAGGCC 3393

Qy     1121 ThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArg 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3394 ACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGATAAAGCCAGG 3453

Qy     1141 ThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeu 1160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3454 ACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTG 3513

Qy     1161 IleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArg 1180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3514 ATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAGCTCCTGAGA 3573

Qy     1181 AsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1195
              |||||||||||||||||||||||||||||||||||||||||||||
Db     3574 AATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3618
```

```
RESULT 2
AX339031
LOCUS       AX339031                3702 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 13 from Patent WO0194400.
ACCESSION   AX339031
VERSION     AX339031.1  GI:18129123
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 13 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .3702
```

```
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
BASE COUNT     1098 a    708 c     856 g    1039 t       1 others
ORIGIN

Alignment Scores:
Pred. No.:                 0             Length:          3702
Score:                     5988.50       Matches:         1195
Percent Similarity:        97.79%        Conservative:    0
Best Local Similarity:     97.79%        Mismatches:      0
Query Match:               99.61%        Indels:          27
DB:                        6             Gaps:            1

US-09-873-409-6 (1-1195) x AX339031 (1-3702)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       34 ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 93

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153

Qy       41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      154 TCTTTCTTCAGGTTGACCCTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGT 213

Qy       61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      214 TACATACAGATTTCCTTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAA 273

Qy       81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      274 CAGTTTTTTCATTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333

Qy      101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      334 GAACTTAACACTCGCATGACAGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATT 393

Qy      121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAG 453

Qy      141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      454 GGCTGGAAACTCACCCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA 513

Qy      161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573

Qy      181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAG 633

Qy      201 GluLysGluLeuGln--------------------------------------------- 205
            |||||||||||||||
Db      634 GAGAAAGAACTTCAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTTGGTTTTATTTT 693

Qy      206 ------------------------------------ArgTyrThrGlnAsnLeuLysAsp 213
                                                ||||||||||||||||||||||||
Db      694 CCCCAGTGGCTACTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGAATCTCAAAGAT 753

Qy      214 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 233
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      754 GCAAAGGATTTTGGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTAC 813

Qy      234 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 253
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      814 TTCTTTATGAATGGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAAT 873

Qy      254 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 273
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      874 GGAGAACCTGGATATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGC 933

Qy      274 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 293
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      934 AGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCT 993

Qy      294 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 313
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      994 GCCTTTCATATTTTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCT 1053

Qy      314 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 333
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1054 GGATATAAACCTGAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTAT 1113

Qy      334 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 353
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1114 CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAG 1173

Qy      354 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 373
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1174 ACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG 1233

Qy      374 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 393
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1234 AGGTTATATGATCCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1293

Qy      394 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 413
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1294 AATGTGCGGCATTATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGG 1353

Qy      414 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 433
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1354 ACCACCATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1413

Qy      434 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 453
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1414 AGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAAT 1473

Qy      454 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 473
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1474 ACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATT 1533

Qy      474 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 493
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1534 GCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTG 1593

Qy      494 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThr 513
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1594 GATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTGGAGAAGGCGAGCAAAGGTCGGACT 1653

Qy      514 ThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeu 533
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1654 ACAATCGTGGTAGCACACCGACTTTCTACTATTCGAAGTGCAGATTTGATTGTGACCCTA 1713

Qy      534 LysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeu 553
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1714 AAGGATGGAATGCTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTA 1773

Qy      554 TyrTyrSerLeuValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMet 573
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1774 TATTATTCACTTGTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATG 1833

Qy      574 ThrTyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLys 593
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1834 ACATATTCTACTGAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAG 1893

Qy      594 SerAspPheIleAspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGlu 613
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1894 TCAGACTTCATTGACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAA 1953

Qy      614 ValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGly 633
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1954 GTCTCTCTATTAAAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGG 2013
```

```
Qy     634 ThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLys 653
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2014 ACATTGGCTTCTGTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAA 2073

Qy     654 IleIleThrMetPheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyr 673
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2074 ATTATAACCATGTTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTAT 2133

Qy     674 SerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeu 693
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2134 TCCATGATATTCGTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTA 2193

Qy     694 PheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAla 713
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2194 TTTTACGGCAGAGCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCC 2253

Qy     714 MetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThr 733
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2254 ATGTTATATCAGGATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACA 2313

Qy     734 ThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyVal 753
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2314 ACAATATTAGCCATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTC 2373

Qy     754 LeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrp 773
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2374 TTAACACAAAATGCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGG 2433

Qy     774 GluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGlu 793
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2434 GAGATGACATTCCTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAA 2493

Qy     794 ThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLys 813
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2494 ACCGCAGCAATGACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAG 2553

Qy     814 IleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAla 833
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2554 ATAGCAACTGAAGCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCC 2613

Qy     834 PheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAla 853
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2614 TTCGAGCAAATGTATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCA 2673

Qy     854 GlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAla 873
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2674 CAGATTATTGGAAGCTGTTATGCATTCAGCCATGCCTTTATATATTTTGCCTATGCAGCA 2733

Qy     874 GlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPhe 893
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2734 GGGTTTCGATTTGGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTC 2793

Qy     894 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 913
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2794 ATAGTTTTTACTGCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCT 2853

Qy     914 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 933
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2854 CCTGAATATTCCAAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAA 2913

Qy     934 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 953
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2914 CCAAATATAGACAGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTA 2973

Qy     954 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 973
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2974 GAGTTTCGAGAAGTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGC 3033

Qy     974 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 993
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3034 TTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGG 3093

Qy     994 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 1013
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3094 AAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTG 3153

Qy    1014 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 1033
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3154 TTTGATGGTGTGGATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATC 3213

Qy    1034 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 1053
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3214 GTTCCTCAAGAGCCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGAC 3273

Qy    1054 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 1073
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3274 AACAGCCGTGTGGTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCAT 3333

Qy    1074 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 1093
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3334 TCTTTTATTGAAGGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAG 3393

Qy    1094 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 1113
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3394 CTTTCTGGCGGCCAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAA 3453

Qy    1114 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 1133
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3454 ATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAG 3513

Qy    1134 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 1153
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3514 CATGCCCTTGATAAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCT 3573

Qy    1154 AlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGly 1173
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3574 GCAATTCAGAACGCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGA 3633

Qy    1174 ThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1193
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3634 ACTCATCAAGAGCTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCA 3693

Qy    1194 ValGln 1195
           ||||||
Db    3694 GTGCAG 3699
```

```
RESULT 3
AX478104
LOCUS       AX478104                3699 bp    DNA     linear   PAT 12-AUG-2002
DEFINITION  Sequence 34 from Patent WO0240541.
ACCESSION   AX478104
VERSION     AX478104.1  GI:22217064
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Tang,Y.T., Yue,H., Nguyen,D.B., Hafalia,A.J., Elliott,V.S., Lu,Y.,
            Walia,N.K., Yao,M.G., Baughn,M.R., Gandhi,A.R., Ding,L.,
            Sanjanwala,M., Ramkumar,J., Arvizu,C., Gietzen,K.J., Lal,P.G.,
            Azimzai,Y., Khan,F.A., Thangavelu,K., Thornton,M., Lu,D.A.,
            Tribouley,C.M., Warren,B.A., Ison,C.H., Das,D., Raumann,B.E.,
            Policky,J.L. and Kearney,L.
  TITLE     Transporters and ion channels
  JOURNAL   Patent: WO 0240541-A 34 23-MAY-2002;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..3699
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 7472030CB1"
BASE COUNT     1116 a    707 c    860 g   1016 t
ORIGIN
```

```
Alignment Scores:
Pred. No.:              0            Length:        3699
Score:                  5760.50      Matches:       1156
Percent Similarity:     95.87%       Conservative:  4
Best Local Similarity:  95.54%       Mismatches:    9
Query Match:            95.82%       Indels:        41
DB:                     6            Gaps:          3

US-09-873-409-6 (1-1195) x AX478104 (1-3699)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||||||||||||||||||||   ||||||   |||||||||||||||||||||||||||
Db     145 ATGATCCTGGGTATACTGACATCACTGTTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 204

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
              :::||||||||||||||||||||||||||||||||||||||||||   |||||||||
Db     205 TGTATAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCACACTAACACAACA 264

Qy      41 SerPhePheArg------------------------------------------LeuThr 46
           ::::::                                                ||||||
Db     265 AATTATCAGAACTGTACTCAGTCTCAAGAGAAGCTGAATGAAGATATGACTCTGTTGACC 324

Qy      47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     325 CTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGTTACATACAGATTTCCTTG 384

Qy      67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     385 TGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAACAGTTTTTCATTCAGTT 444

Qy      87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     445 TTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGTGAACTTAACACTCGCATG 504

Qy     107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
           |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     505 ACAGATGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATTGCTCTGTTGTTTCAA 564

Qy     126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     565 AACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAGGGCTGGAAACTCACC 624

Qy     146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     625 CTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCAGCATGTTCTAGGATG 684

Qy     166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     685 GTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCAGAA 744

Qy     186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||-|||||
Db     745 GAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAGGAGAAAGAACTTCAA 804

Qy     206 ArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSer 225
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     805 AGGTATACACAGAATCTCAAAGATGCAAAGGATTTTGGCATAAAAAGGACTATAGCTTCA 864

Qy     226 LysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrp 245
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     865 AAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAATGGAACCTATGGACTTGCTTTTTGG 924

Qy     246 TyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAla 265
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     925 TATGGAACCTCCTTGATTCTTAATGGAGAACCTGGATATACCATCGGGACTGTTCTTGCT 984

Qy     266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     985 GTTTTCTTTAGTGTAATCCATAGTAGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAA 1044

Qy     286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1045 ACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTTATTGATAAGAAACCC 1104

Qy     306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
           ||||||      ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1105 AGTATAGGTAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGAA 1164

Qy     326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1165 TTTAAAAATGTTTCTTTCAATTATCCATCAAGGCCATCTATCAAGATTCTGAAAGGTCTG 1224

Qy     346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
           ||||||      ||||||||||||||||||||||||||||||||||||||||||||||||
Db    1225 AATCTCGGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAG 1284

Qy     366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1285 AGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGTG 1344

Qy     386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1345 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTGGTT 1404

Qy     406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1405 AGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGAGAT 1464

Qy     426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1465 GATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATC 1524

Qy     446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1525 ATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGA 1584

Qy     466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1585 GGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATT 1644

Qy     486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeu 505
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1645 TTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTG 1704

Qy     506 GluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArg 525
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1705 GAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTAGCACACCGACTTTCTACTATTCGA 1764

Qy     526 SerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAla 545
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1765 AGTGCAGATTTGATTGTGACCCTAAAGGATGGAATGCTGGCGGAGAAAGGAGCGCATGCT 1824

Qy     546 GluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLys 565
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1825 GAACTAATGGCAAAACGAGGTCTATATTATTCACTTGTGATGTCACAGGATATTAAAAAA 1884

Qy     566 AlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPro 585
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1885 GCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCACTTCCT 1944

Qy     586 LeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGln 605
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1945 CTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCCACCCAA 2004

Qy     606 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 625
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2005 TCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAACAAGCCT 2064

Qy     626 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 645
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2065 GAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCTGTTCTAAATGGAACTGTTCATCCA 2124

Qy     646 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 665
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2125 GTATTTTCCATCATCTTTGCAAAAATTATAACCATGTTTGGAAATAATGATAAAACCACA 2184
```

```
Qy    666 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 685
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2185 TTAAAGCATGATGCAGAAATTTATTCCATGATATTCGTCATTTTGGGTGTTATTTGCTTT 2244

Qy    686 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 705
          |||||||||||||||
Db   2245 GTCAGTTATTTCATG--------------------------------------------- 2259

Qy    706 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLys 725
                                           |||||||||||||||||||||||||||
Db   2260 ---------------------------------CAGGATATTGCCTGGTTTGATGAAAAG 2286

Qy    726 GluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 745
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2287 GAAAACAGCACAGGAGGCTTGACAACAATATTAGCCATAGATATAGCACAAATTCAAGGA 2346

Qy    746 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 765
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2347 GCAACAGGTTCCAGGATTGGCGTCTTAACACAAAATGCAACTAACATGGGACTTTCAGTT 2406

Qy    766 IleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVal 785
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2407 ATCATTTCCTTTATATATGGATGGGAGATGACATTCCTGATTCTGAGTATTGCTCCAGTA 2466

Qy    786 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 805
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2467 CTTGCCGTGACAGGAATGATTGAAACCGCAGCAATGACTGGATTTGCCAACAAAGATAAG 2526

Qy    806 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIle 825
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2527 CAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAGAATATACGTACTATA 2586

Qy    826 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGln 845
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2587 GTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATGTATGAAGAGATGCTTCAGACTCAA 2646

Qy    846 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla 865
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2647 CACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTATGCATTCAGCCATGCC 2706

Qy    866 PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGly 885
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2707 TTTATATATTTTGCCTATGCGGCAGGGTTTCGATTTGGAGCCTATTTAATTCAAGCTGGA 2766

Qy    886 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAla 905
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2767 CGAATGACCCCAGAGGGCATGTTCATAGTTTTTACTGCAATTGCATATGGAGCTATGGCC 2826

Qy    906 IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHis 925
          ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2827 ATCGGAGAAACGCTCGTTTTGGCTCCTGAATATTCCAAAGCCAAATCGGGGGCTGCGCAT 2886

Qy    926 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLys 945
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2887 CTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGACAGCCGCAGTCAAGAAGGGAAAAAG 2946

Qy    946 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 965
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2947 CCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTCTTCTATCCATGTCGC 3006

Qy    966 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 985
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3007 CCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGCA 3066

Qy    986 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr 1005
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3067 TTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTAT 3126

Qy   1006 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 1025
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3127 GACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTGGATGCAAAAGAATTGAATGTACAG 3186

Qy   1026 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 1045
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3187 TGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAGCCTGTGCTCTTCAACTGCAGCATT 3246

Qy   1046 AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu 1065
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3247 GCTGAGAACATCGCCTATGGTGACAACAGCCGTGTGGTGCCATTAGATGAGATCAAAGAA 3306

Qy   1066 AlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThr 1085
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3307 GCCGCAAATGCAGCAAATATCCATTCTTTTATTGAAGGTCTCCCTGAGAAATACAACACA 3366

Qy   1086 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAla 1105
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3367 CAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCTATTGCA 3426

Qy   1106 ArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsp 1125
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3427 AGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGAT 3486

Qy   1126 AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys 1145
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3487 AATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGATAAAGCCAGGACGGGAAGGACATGC 3546

Qy   1146 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 1165
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3547 CTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGGTTCTGCAC 3606

Qy   1166 AsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyr 1185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3607 AATGGAAAGATAAAGGAACAAGGAACTCATCAAGAGCTCCTGAGAAATCGAGACATATAT 3666

Qy   1186 PheLysLeuValAsnAlaGlnSerValGln 1195
          ||||||||||||||||||||||||||||||
Db   3667 TTTAAGTTAGTGAATGCACAGTCAGTGCAG 3696
```

```
RESULT 4
AX339030
LOCUS       AX339030                3177 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 12 from Patent WO0194400.
ACCESSION   AX339030
VERSION     AX339030.1  GI:18129122
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 12 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..3177
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      970 a    601 c    735 g    870 t     1 others
ORIGIN

Alignment Scores:
Pred. No.:              0              Length:         3177
Score:                  5159.50        Matches:        1031
Percent Similarity:     97.45%         Conservative:   0
Best Local Similarity:  97.45%         Mismatches:     0
Query Match:            85.82%         Indels:         27
DB:                     6              Gaps:           1

US-09-873-409-6 (1-1195) x AX339030 (1-3177)

Qy    165 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 184
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 ATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCA 60
```

```
Qy    185 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 204
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAGGAGAAAGAACTT 120

Qy    205 Gln--------------------------------------------------------- 205
          |||
Db    121 CAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTTGGTTTTATTTCCCCAGTGGCTA 180

Qy    206 ------------------------ArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 217
                                  ||||||||||||||||||||||||||||||||||||
Db    181 CTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGAATCTCAAAGATGCAAAGGATTTT 240

Qy    218 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 237
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAAT 300

Qy    238 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 257
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGA 360

Qy    258 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 277
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATT 420

Qy    278 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 297
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATT 480

Qy    298 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 317
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCT 540

Qy    318 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 337
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCA 600

Qy    338 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 357
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTG 660

Qy    358 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 377
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 720

Qy    378 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 397
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCAT 780

Qy    398 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 417
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGT 840

Qy    418 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 437
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGG 900

Qy    438 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 457
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGG 960

Qy    458 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 477
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTA 1020

Qy    478 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 497
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGC 1080

Qy    498 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 517
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 AAGTCAGCTGTTCAAGCTGCACTGGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTA 1140

Qy    518 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 537
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 GCACACCGACTTTCTACTATTCGAAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATG 1200

Qy    538 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 557
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 CTGGCGGAGAAAGGAGCACATGCTGAACTAATGGCAAAACGAGGTCTATATTATTCACTT 1260

Qy    558 ValMetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThr 577
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 GTGATGTCACAGGATATTAAAAAAGCTGATGAACAGATGGAGTCAATGACATATTCTACT 1320

Qy    578 GluArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle 597
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GAAAGAAAGACCAACTCACTTCCTCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATT 1380

Qy    598 AspLysAlaGluGluSerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeu 617
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 GACAAGGCTGAGGAATCCACCCAATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTA 1440

Qy    618 LysIleLeuLysLeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSer 637
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 AAAATTTTAAAGTTAAACAAGCCTGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCT 1500

Qy    638 ValLeuAsnGlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMet 657
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 GTTCTAAATGGAACTGTTCATCCAGTATTTTCCATCATCTTTGCAAAAATTATAACCATG 1560

Qy    658 PheGlyAsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePhe 677
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 TTTGGAAATAATGATAAAACCACATTAAAGCATGATGCAGAAATTTATTCCATGATATTC 1620

Qy    678 ValIleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArg 697
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 GTCATTTTGGGTGTTATTTGCTTTGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGA 1680

Qy    698 AlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGln 717
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GCAGGGGAAATTTTAACGATGAGATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAG 1740

Qy    718 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 737
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 GATATTGCCTGGTTTGATGAAAAGGAAAACAGCACAGGAGGCTTGACAACAATATTAGCC 1800

Qy    738 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 757
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ATAGATATAGCACAAATTCAAGGAGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAAT 1860

Qy    758 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 777
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GCAACTAACATGGGACTTTCAGTTATCATTTCCTTTATATATGGATGGGAGATGACATTC 1920

Qy    778 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 797
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CTGATTCTGAGTATTGCTCCAGTACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATG 1980

Qy    798 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 817
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 ACTGGATTTGCCAACAAAGATAAGCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAA 2040

Qy    818 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 837
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 GCTTTGGAGAATATACGTACTATAGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATG 2100

Qy    838 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 857
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 TATGAAGAGATGCTTCAGACTCAACACAGAAATACCTCGAAGAAAGCACAGATTATTGGA 2160

Qy    858 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 877
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 AGCTGTTATGCATTCAGCCATGCCTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTT 2220

Qy    878 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 897
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db   2221 GGAGCCTATTTAATTCAAGCTGGACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACT 2280

Qy    898 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 917
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 GCAATTGCATATGGAGCTATGGCCATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCC 2340

Qy    918 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 937
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 AAAGCCAAATCGGGGGCTGCGCATCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGAC 2400

Qy    938 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 957
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 AGCCGCAGTCAAGAAGGGAAAAAGCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAA 2460

Qy    958 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSer 977
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 GTCTCTTTCTTCTATCCATGTCGCCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGT 2520

Qy    978 IleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSer 997
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 ATTGAGCGAGGAAAGACAGTAGCATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCT 2580

Qy    998 ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyVal 1017
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 GTTCAACTTCTGCAGAGACTTTATGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTG 2640

Qy   1018 AspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGlu 1037
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 GATGCAAAAGAATTGAATGTACAGTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAG 2700

Qy   1038 ProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgVal 1057
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 CCTGTGCTCTTCAACTGCAGCATTGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTG 2760

Qy   1058 ValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGlu 1077
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 GTGCCATTAGATGAGATCAAAGAAGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAA 2820

Qy   1078 GlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGly 1097
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 GGTCTCCCTGAGAAATACAACACACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGC 2880

Qy   1098 GlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeu 1117
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 CAGAAACAAAGACTAGCTATTGCAAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTG 2940

Qy   1118 AspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAsp 1137
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 GATGAGGCCACTTCAGCCCTCGATAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGAT 3000

Qy   1138 LysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsn 1157
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 AAAGCCAGGACGGGAAGGACATGCCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAAC 3060

Qy   1158 AlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGlu 1177
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 GCAGATTTGATAGTGGTTCTGCACAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAG 3120

Qy   1178 LeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerValGln 1195
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 CTCCTGAGAAATCGAGACATATATTTTAAGTTAGTGAATGCACAGTCAGTGCAG 3174

RESULT 5
AX339028
LOCUS       AX339028               2856 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 10 from Patent WO0194400.
ACCESSION   AX339028
VERSION     AX339028.1  GI:18129120
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 10 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..2856
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      888 a    540 c    652 g    776 t
ORIGIN

Alignment Scores:
Pred. No.:              8.1e-288      Length:         2856
Score:                  4561.50       Matches:        912
Percent Similarity:     97.96%        Conservative:   0
Best Local Similarity:  97.96%        Mismatches:     1
Query Match:            75.87%        Indels:         18
DB:                     6             Gaps:           1

US-09-873-409-6 (1-1195) x AX339028 (1-2856)

Qy    266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHis-PheGl 285
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 
Db    114 GTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATTGGAGCAGCAGTCCCTCATTATT-- 171

Qy    285 uThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPr 305
                                                           |||||||||||
Db    172 -------------------------------------------------GATAAGAAACC 182

Qy    305 oSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGl 325
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    183 CAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGA 242

Qy    325 uPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLe 345
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    243 ATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCT 302

Qy    345 uAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLy 365
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    303 GAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGAA 362

Qy    365 sSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVa 385
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    363 GAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGT 422

Qy    385 lAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVa 405
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    423 GGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTGGT 482

Qy    405 lSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAs 425
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    483 TAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGAGA 542

Qy    425 pAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIl 445
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    543 TGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTTAT 602

Qy    445 eMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGl 465
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    603 CATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGG 662

Qy    465 yGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIl 485
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    663 AGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGAT 722

Qy    485 eLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLe 505
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    723 TTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACT 782

Qy    505 uGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleAr 525
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     783 GGAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTAGCACACCGACTTTCTACTATTCG 842

Qy     525 gSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAl 545
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     843 AAGTGCAGATTTGATTGTGACCCTAAAGGATGGAATGCTGGCGGAGAAAGGAGCACATGC 902

Qy     545 aGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLy 565
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     903 TGAACTAATGGCAAAACGAGGTCTATATTATTCACTTGTGATGTCACAGGATATTAAAAA 962

Qy     565 sAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPr 585
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     963 AGCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCACTTCC 1022

Qy     585 oLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGl 605
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1023 TCTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCCACCCA 1082

Qy     605 nSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPr 625
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1083 ATCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAACAAGCC 1142

Qy     625 oGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPr 645
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1143 TGAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCTGTTCTAAATGGAACTGTTCATCC 1202

Qy     645 oValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrTh 665
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1203 AGTATTTTCCATCATCTTTGCAAAAATTATAACCATGTTTGGAAATAATGATAAAACCAC 1262

Qy     665 rLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPh 685
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1263 ATTAAAGCATGATGCAGAAATTTATTCCATGATATTCGTCATTTTGGGTGTTATTTGCTT 1322

Qy     685 eValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetAr 705
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1323 TGTCAGTTATTTCATGCAGGGATTATTTTACGGCAGAGCAGGGGAAATTTTAACGATGAG 1382

Qy     705 gLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLy 725
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1383 ATTAAGACACTTGGCCTTCAAAGCCATGTTATATCAGGATATTGCCTGGTTTGATGAAAA 1442

Qy     725 sGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGl 745
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1443 GGAAAACAGCACAGGAGGCTTGACAACAATATTAGCCATAGATATAGCACAAATTCAAGG 1502

Qy     745 yAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVa 765
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1503 AGCAACAGGTTCCAGGATTGGCGTCTTAACACAAAATGCAACTAACATGGGACTTTCAGT 1562

Qy     765 lIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVa 785
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1563 TATCATTTCCTTTATATATGGATGGGAGATGACATTCCTGATTCTGAGTATTGCTCCAGT 1622

Qy     785 lLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLy 805
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1623 ACTTGCCGTGACAGGAATGATTGAAACCGCAGCAATGACTGGATTTGCCAACAAAGATAA 1682

Qy     805 sGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIl 825
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1683 GCAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAGAATATACGTACTAT 1742

Qy     825 eValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGl 845
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1743 AGTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATGTATGAAGAGATGCTTCAGACTCA 1802

Qy     845 nHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAl 865
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1803 ACACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTATGCATTCAGCCATGC 1862

Qy     865 aPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGl 885
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1863 CTTTATATATTTTGCCTATGCAGCAGGGTTTCGATTTGGAGCCTATTTAATTCAAGCTGG 1922

Qy     885 yArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAl 905
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1923 ACGAATGACCCCAGAGGGCATGTTCATAGTTTTTACTGCAATTGCATATGGAGCTATGGC 1982

Qy     905 aIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHi 925
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1983 CATCGGAAAAACGCTCGTTTTGGCTCCTGAATATTCCAAAGCCAAATCGGGGGCTGCGCA 2042

Qy     925 sLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLy 945
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2043 TCTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGACAGCCGCAGTCAAGAAGGGAAAAA 2102

Qy     945 sProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysAr 965
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2103 GCCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTCTTCTATCCATGTCG 2162

Qy     965 gProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAl 985
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2163 CCCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGC 2222

Qy     985 aPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTy 1005
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2223 ATTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTA 2282

Qy    1005 rAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGl 1025
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2283 TGACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTGGATGCAAAAGAATTGAATGTACA 2342

Qy    1025 nTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIl 1045
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2343 GTGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAGCCTGTGCTCTTCAACTGCAGCAT 2402

Qy    1045 eAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGl 1065
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2403 TGCTGAGAACATCGCCTATGGTGACAACAGCCGTGTGGTGCCATTAGATGAGATCAAAGA 2462

Qy    1065 uAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnTh 1085
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2463 AGCCGCAAATGCAGCAAATATCCATTCTTTTATTGAAGGTCTCCCTGAGAAATACAACAC 2522

Qy    1085 rGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAl 1105
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2523 ACAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCTATTGC 2582

Qy    1105 aArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAs 1125
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2583 AAGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGA 2642

Qy    1125 pAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCy 1145
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2643 TAATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGATAAAGCCAGGACGGGAAGGACATG 2702

Qy    1145 sLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHi 1165
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2703 CCTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGGTTCTGCA 2762

Qy    1165 sAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTy 1185
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2763 CAATGGAAAGATAAAGGAACAAGGAACTCATCAAGAGCTCCTGAGAAATCGAGACATATA 2822

Qy    1185 rPheLysLeuValAsnAlaGlnSerValGln 1195
           ||||||||||||||||||||||||||||||
Db    2823 TTTTAAGTTAGTGAATGCACAGTCAGTGCAG 2853

RESULT 6
CRUPGPI
LOCUS       CRUPGPI                 4296 bp    mRNA    linear   ROD 22-JUN-1995
DEFINITION  C.griseus P-glycoprotein (isoform class I) mRNA, complete cds.
ACCESSION   M60040
VERSION     M60040.1  GI:191164
KEYWORDS    multidrug resistance; p-glycoprotein; transmembrane protein.
SOURCE      C.griseus adult liver and ovary, cDNA to mRNA.
```

```
  ORGANISM  Cricetulus griseus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
            Cricetulus.
REFERENCE   1  (bases 1 to 4296)
  AUTHORS   Endicott,J.A., Sarangi,F. and Ling,V.
  TITLE     Complete cDNA sequences encoding the Chinese hamster P-glycoprotein
            gene family
  JOURNAL   DNA Seq. 2 (2), 89-101 (1991)
  MEDLINE   92135896
   PUBMED   1685679
FEATURES             Location/Qualifiers
     source          1..4296
                     /organism="Cricetulus griseus"
                     /db_xref="taxon:10029"
                     /tissue_type="liver and ovary"
                     /dev_stage="adult"
     CDS             103..3933
                     /note="putative"
                     /codon_start=1
                     /product="p-glycoprotein isoform I"
                     /protein_id="AAA68883.1"
                     /db_xref="GI:191165"
                     /translation="MEFEEDFSGRKDKNFLKMGRKSKKEKKEKKPVVSVFTMFRYAGW
                     LDRLYMLVGTLAAIIHGVALPLMMLVFGDMTDSFASVGNIPTNATNNATQVNASDIFG
                     KLEEEMTTYAYYYTGIGAGVLIVAYIQVSFWCLAAGRQIHKIRQKFFHAIMNQEIGWF
                     DVHDVGELNTRLTDDVSKINEGIGDKIGMFFQAMATFFGGFIIGFTRGWKLTLVILAI
                     SPVLGLSAGIWAKILSSFTDKELQAYAKAGAVAEEVLAAIRTVIAFGGQKKELERYNN
                     NLEEAKRLGIKKAITANISMGAAFLLIYASYALAFWYGTSLVISKEYSIGQVLTVFFA
                     VLIGAFSIGQASPNIEAFANARGAAYEIFNIIDNKPSIDSFSKNGYKPDNIKGNLEFK
                     NIHFSYPSRKDVQILKGLNLKVQSGQTVALVGNSGCGKSTTVQLLQRLYDPTEGVVSI
                     DGQDIRTINVRYLREIIGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKEANAYD
                     FIMKLPHKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTESEAVV
                     QAALDKAREGRTTIVIAHRLSTVRNADIIAGFDGGVIVEQGNHEELMREKGIYFKLVM
                     TQTAGNEIELGNEVGESKNEIDNLDMSSKDSASSLIRRRSTRRSIRGPHDQDRKLSTK
                     EALDEDVPPISFWRILKLNSSEWPYFVVGIFCAIVNGALQPAFSIIFSKVVGVFTRNT
                     DDETKRHDSNLFSLLFLILGVISFITFFLQGFTFGKAGEILTKRLRYMVFKSMLRQDV
                     SWFDNPKNTTGALTTRLANDAGQVKGATGARLAVITQNIANLGTGIIISLIYGWQLTL
                     LLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGSGKIATEAIENFRTVVSLTREQKFE
                     NMYAQSLQIPYRNALKKAHVFGITFSFTQAMMYFSYAACFRFGAYLVARELMTFENVL
                     LVFSAIVFGAMAVGQVSSFAPDYAKAKVSASHIIMIIEKVPSIDSYSTGGLKPNTLEG
                     NVKFNEVVFNYPTRPDIPVLQGLNLEVKKGQTLALVGSSGCGKSTVVQLLERFYDPMA
                     GTVFLDGKEVNQLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSQDEIERAA
                     KEANIHQFIESLPDKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEATSALD
                     TESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVIQNGKVKEHGTHQQLLAQKG
                     IYFSMVSVQAGAKR"
BASE COUNT     1271 a    892 c   1029 g   1104 t
ORIGIN

Alignment Scores:
Pred. No.:              2.61e-223     Length:         4296
Score:                  3574.00       Matches:        679
Percent Similarity:     75.37%        Conservative:   248
Best Local Similarity:  55.20%        Mismatches:     259
Query Match:            59.45%        Indels:         44
DB:                     10            Gaps:           8

US-09-873-409-6 (1-1195) x CRUPGPI (1-4296)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            |||::::::|||  ||||||:::::::::::|||          |||||||||||   |||
Db      250 ATGCTGGTGGGGACTCTCGCAGCTATTATCCATGGAGTTGCACTCCCGCTTATGATGCTG 309

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsn------ 38
            |||    |||:::|||:::|||:::          |||         :::  |||||
Db      310 GTGTTTGGAGACATGACAGACAGCTTTGCAAGTGTAGGAAACATCCCAACAAATGCTACT 369

Qy       38 ------------------------------------------------------------ 38

Db      370 AATAATGCCACTCAAGTCAATGCTTCAGACATCTTTGGCAAACTGGAGGAAGAAATGACC 429

Qy       39 ThrTyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIle 58
            ||||||::::::                  ||||||   |||||||||         ||||||
Db      430 ACGTACGCCTAT--------------TACTACACAGGGATTGGTGCTGGTGTGCTTATA 474

Qy       59 PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIle 78
               |||||||||:::|||   |||   :::   |||   ||||||         :::|||
Db      475 GTTGCCTACATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATT 534

Qy       79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp 98
            |||::::::|||||||||:::::::::      |||:::|||||||||||||||      |||
Db      535 AGACAGAAGTTTTTCATGCTATAATGAATCAGGAGATAGGCTGGTTTGACGTGCACGAC 594

Qy       99 IleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly 117
            :::|||||||||||||||||||:::|||   |||:::   ||||||::::::||||||||
Db      595 GTTGGGGAGCTCAACACCCGGCTCACAGATGATGTCTCCAAAATTAACGAAGGAATTGGT 654

Qy      118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
            |||||||||   :::   ||||||   |||:::||||||         |||      :::|||
Db      655 GACAAAATTGGAATGTTCTTTCAGGCAATGGCAACATTTTTTGGTGGTTTTATAATAGGA 714

Qy      138 LeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
               :::||||||||||||||||||||||   |||:::   ||||||::::::
Db      715 TTTACTCGAGGCTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTG 774

Qy      158 SerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr 177
            ||||||      :::::::::::   |||   |||   |||||||||   ||||||
Db      775 TCAGCTGGTATTTGGGCAAAGATATTATCTTCATTTACTGATAAAGAACTTCAGGCATAT 834

Qy      178 SerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197
            :::||||||||||||||||||||||||||||||||||:::::|||||||||||||||||
Db      835 GCAAAAGCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCATTT 894

Qy      198 ArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 217
               |||:::|||||||||:::||||||         ||||||::::::||||||
Db      895 GGAGGACAAAAGAAAGAACTTGAAAGGTACAACAACAATTTAGAAGAAGCCAAAAGACTT 954

Qy      218 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 237
            |||||||||:::   |||   :::   :::|||:::||||||   :::      :::
Db      955 GGGATAAAGAAAGCGATCACAGCCAACATTTCCATGGGCGCAGCTTTCCTTCTTATCTAT 1014

Qy      238 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 257
               :::|||   ||||||||||||||||||||||||||||:::::::::   |||
Db     1015 GCATCATATGCTCTGGCCTTCTGGTATGGAACGTCCTTGGTCATCTCAAAAGAA------ 1068

Qy      258 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 277
            |||:::||||   ||||||   ||||||||||:::|||:::      ::::::   |||
Db     1069 TATTCTATTGGACAAGTGCTCACTGTCTTCTTTGCTGTATTAATTGGGGCATTCAGTATT 1128

Qy      278 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 297
            |||   |||   |||:::   |||   ||||||   ||||||||||||||:::   |||
Db     1129 GGACAGGCATCTCCAAATATCGAAGCCTTTGCCAATGCAAGAGGAGCAGCCTATGAAATC 1188

Qy      298 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 317
            |||   :::||||||   ||||||||||||||:::||||||      ||||||||||||
Db     1189 TTCAACATAATTGATAATAAGCCCAGTATTGACAGCTTCTCAAAGAATGGGTACAAACCA 1248

Qy      318 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 337
            ::::::|||:::|||   :::||||||||||||:::   |||:::||||||||||||
Db     1249 GACAACATTAAAGGAAATTTGGAATTCAAAAATATTCACTTCAGTTACCCATCTCGAAAA 1308

Qy      338 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 357
               ::::::|||||||||||||||||||||:::::::::||||||:::|||||||||||||
Db     1309 GACGTTCAGATCTTGAAGGGCCTCAACCTGAAGGTGCAGAGCGGACAGACAGTGGCCCTG 1368

Qy      358 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 377
            ||||||      :::|||   ||||||||||||   |||||||||||||||||||||
Db     1369 GTTGGCAACAGTGGCTGTGGGAAAAGCACCACTGTCCAGCTGCTGCAGAGGCTCTACGAC 1428

Qy      378 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 397
            |||   :::|||   :::   :::|||         ||||||||      :::|||||||:::
Db     1429 CCCACAGAGGGCGTGGTCAGTATCGACGGTCAGGACATCAGGACCATCAATGTGAGGTAT 1488

Qy      398 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 417
               |||:::   ||||||||||||||||||||||||||||||   |||||||||:::
Db     1489 CTGCGGGAAATCATTGGGGTCGTGAGTCAGGAACCTGTGTTGTTTGCCACCACTATCGCT 1548
```

```
Qy    418 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 437
Db   1549 GAAAACATTCGCTATGGCCGAGAAAATGTCACTATGGACGAGATCGAGAAAGCTGTCAAG 1608

Qy    438 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 457
Db   1609 GAAGCCAATGCCTATGACTTCATCATGAAACTGCCCCATAAATTTGACACTCTTGTTGGT 1668

Qy    458 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 477
Db   1669 GAGAGAGGAGCACAGCTGAGTGGAGGACAGAAACAGAGAATCGCCATTGCTCGTGCCCTG 1728

Qy    478 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 497
Db   1729 GTCCGCAACCCCAAGATCCTTTTGTTGGATGAGGCGACATCAGCCTTGGACACAGAAAGT 1788

Qy    498 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 517
Db   1789 GAAGCCGTGGTTCAGGCTGCTCTGGATAAGGCTAGAGAAGGCCGGACTACCATTGTGATA 1848

Qy    518 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 537
Db   1849 GCCCACCGCTTGTCTACAGTTCGAAATGCTGACATCATAGCTGGGTTTGATGGTGGTGTC 1908

Qy    538 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 557
Db   1909 ATTGTGGAGCAAGGAAATCATGAAGAGCTCATGAGAGAGAAGGGCATTTACTTCAAACTT 1968

Qy    558 ValMetSerGln------------------------------AspIleLysLysAlaAsp 567
Db   1969 GTCATGACACAGACAGCAGGAAATGAAATTGAATTAGGAAATGAAGTTGGTGAGTCTAAA 2028

Qy    568 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu------Pro 585
Db   2029 AATGAAATTGATAATTTAGACATGTCTTCAAAAGATTCAGCATCCAGTCTAATCAGAAGG 2088

Qy    586 LeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr--- 604
Db   2089 AGATCAACTCGCAGAAGTATCCGTGGACCACATGACCAAGACAGGAAGCTTAGTACCAAA 2148

Qy    605 GlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLys 624
Db   2149 GAGGCCTTGGATGAAGATGTACCTCCAATTTCCTTTTGGCGTATCCTAAAGTTGAATTCA 2208

Qy    625 ProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHis 644
Db   2209 TCTGAATGGCCATATTTTGTGGTTGGTATATTCTGTGCCATAGTAAATGGAGCCTTGCAA 2268

Qy    645 ProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLys 663
Db   2269 CCAGCATTCTCAATAATATTTTCTAAGGTTGTAGGGGTTTTCACAAGAAATACTGATGAT 2328

Qy    664 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 683
Db   2329 GAAACCAAACGACATGATAGCAACTTGTTTTCCTTATTGTTTCTGATCCTTGGGGTCATT 2388

Qy    684 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 703
Db   2389 TCTTTTATTACATTTTTTCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACC 2448

Qy    704 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 723
Db   2449 AAGCGACTCCGATACATGGTTTTCAAATCCATGCTGAGACAGGATGTGAGCTGGTTTGAT 2508

Qy    724 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 743
Db   2509 AACCCTAAAAACACCACTGGAGCATTGACCACCAGGCTCGCCAACGATGCTGGTCAAGTT 2568

Qy    744 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 763
Db   2569 AAAGGGGCTACAGGAGCCAGACTTGCTGTCATTACCCAGAATATAGCAAATCTTGGGACA 2628

Qy    764 SerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAla 783
Db   2629 GGAATCATCATATCCCTAATCTATGGCTGGCAGTTGACACTTCTACTCTTAGCAATTGTC 2688

Qy    784 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 803
Db   2689 CCCATCATTGCAATAGCAGGAGTGGTTGAGATGAAAATGTTGTCTGGACAAGCACTAAAA 2748

Qy    804 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 823
Db   2749 GATAAGAAGGAGCTAGAAGGTTCTGGGAAGATTGCTACTGAAGCAATAGAGAACTTCCGC 2808

Qy    824 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 843
Db   2809 ACTGTCGTCTCTTTGACTCGGGAGCAGAAGTTTGAAAATATGTATGCCCAGAGCTTGCAG 2868

Qy    844 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 863
Db   2869 ATACCATACAGAAATGCTCTGAAGAAAGCACACGTCTTTGGGATCACATTCTCCTTCACC 2928

Qy    864 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 883
Db   2929 CAGGCCATGATGTATTTTTCGTATGCTGCTTGTTTCCGGTTTGGTGCTTACTTGGTGGCA 2988

Qy    884 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 903
Db   2989 CGGGAATTAATGACATTTGAAAATGTTCTATTAGTATTCTCAGCTATTGTCTTTGGTGCC 3048

Qy    904 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 923
Db   3049 ATGGCAGTTGGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTGTCGGCA 3108

Qy    924 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 943
Db   3109 TCGCATATCATCATGATCATTGAAAAAGTCCCTTCCATTGACAGCTACAGCACCGGAGGC 3168

Qy    944 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 963
Db   3169 CTGAAGCCTAATACATTGGAAGGAAATGTGAAATTTAATGAAGTCGTGTTCAACTACCCC 3228

Qy    964 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 983
Db   3229 ACCCGGCCTGACATCCCAGTGCTTCAGGGGCTGAACCTGGAGGTGAAGAAGGGCCAGACG 3288

Qy    984 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 1003
Db   3289 CTGGCCCTGGTGGGCAGCAGTGGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAGCGG 3348

Qy   1004 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 1023
Db   3349 TTCTACGACCCCATGGCCGGAACAGTGTTTCTAGATGGCAAAGAAGTAAACCAACTAAAT 3408

Qy   1024 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 1043
Db   3409 GTCCAGTGGCTCCGAGCACACCTAGGCATTGTGTCCCAAGAGCCCATCCTGTTTGACTGC 3468

Qy   1044 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 1063
Db   3469 AGCATCGCTGAGAACATCGCTTATGGAGACAACAGCCGGGTCGTGTCCCAGGATGAGATC 3528

Qy   1064 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 1083
Db   3529 GAGAGGGCGGCGAAGGAAGCCAACATCCACCAGTTCATTGAGTCACTGCCTGACAAATAC 3588

Qy   1084 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 1103
Db   3589 AACACCAGAGTAGGAGACAAAGGGACTCAGCTGTCGGGTGGGCAGAAACAACGCATCGCC 3648

Qy   1104 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 1123
Db   3649 ATAGCACGCGCCCTCGTCAGACAGCCTCACATTTTGCTCCTGGATGAAGCAACATCAGCC 3708

Qy   1124 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 1143
```

```
             |||||  :::|||||||||||||||||  |||||||||||||||||  ||||||
Db      3709 CTGGATACAGAGAGTGAAAAGGTTGTCCAGGAAGCTCTGGACAAAGCCAGAGAAGGCCGC 3768

Qy      1144 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 1163
             |||||:::|||:::  ||||||||||||  ||||||||||||||||||||||||
Db      3769 ACCTGCATTGTGATCGCTCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTG 3828

Qy      1164 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 1183
             :::  ||||||||:::|||||  ||||||||||||:::||||||        :::
Db      3829 ATTCAGAATGGCAAGGTCAAGGAGCATGGCACCCACCAGCAGCTGCTGGCACAGAAAGGC 3888

Qy      1184 IleTyrPheLysLeuValAsnAlaGlnSer 1193
             ||||||||   :::|||:::  |||:::
Db      3889 ATCTATTTCTCCATGGTCAGTGTGCAGGCT 3918
```

RESULT 7
CRUPGP1165

```
LOCUS       CRUPGP1165            4304 bp    mRNA    linear   ROD 27-APR-1993
DEFINITION  Chinese hamster p-glycoprotein mRNA (clone ADX165), complete cds.
ACCESSION   M59253
VERSION     M59253.1  GI:191154
KEYWORDS    multidrug resistance glycoprotein; p-glycoprotein.
SOURCE      Chinese hamster, cDNA to mRNA, clone ADX165.
  ORGANISM  Cricetulus sp.
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
            Cricetulus.
REFERENCE   1  (bases 1 to 4304)
  AUTHORS   Devine,S.E., Hussain,A., Davide,J.P. and Melera,P.W.
  TITLE     Full length and alternatively spliced pgp1 transcripts in
            multidrug-resistant Chinese hamster lung cells
  JOURNAL   J. Biol. Chem. 266 (7), 4545-4555 (1991)
  MEDLINE   91154265
   PUBMED   1671863
FEATURES             Location/Qualifiers
     source          1..4304
                     /organism="Cricetulus sp."
                     /db_xref="taxon:10031"
                     /cell_type="fibroblast"
                     /tissue_type="lung"
     gene            1..4304
                     /gene="pgp-1"
     CDS             111..3941
                     /gene="pgp-1"
                     /codon_start=1
                     /product="p-glycoprotein"
                     /protein_id="AAA37004.1"
                     /db_xref="GI:191155"
                     /translation="MEFEEDFSGRKDKNFLKMGRKSKKEKKEKKPVVSVFTMFRYAGW
                     LDRLYMLVGTLAAIIHGVALPLMMLVFGDMTDSFASVGNIPTNATNNATQVNASDIFG
                     KLEEEMTTYAYYYTGIGAGVLIVAYIQVSFWCLAAGRQIHKIRQKFFHAIMNQEIGWF
                     DVHDVGELNTRLTDDVSKINEGIGDKIGMFFQAMATFFGGFIIGFTRGWKLTLVILAI
                     SPVLGLSAGIWAKILSSFTDKELQAYAKAGAVAEEVLAAIRTVIAFGGQKKELERYNN
                     NLEEAKRLGIKKAITANISMGAAFLLIYASYALAFWYGTSLVISKEYSIGQVLTVFFA
                     VLIAPFSIGQASPNIEAFANARGAAYEIFNIIDNKPSIDSFSKNGYKPDNIKGNLEFK
                     NIHFSYPSRKDVQILKGLNLKVQSGQTVALVGNSGCGKSTTVQLLQRLYDPTEGVVSI
                     DGQDIRTINVRYLREIIGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKEANAYD
                     FIMKLPHKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTESEAVV
                     QAALDKAREGRTTIVIAHRLSTVRNADIIAGFDGGVIVEQGNHEELMREKGIYFKLVM
                     TQTAGNEIELGNEVGESKNEIDNLDMSSKDSASSLIRRRSTRRSIRGPHDQDRKLSTK
                     EALDEDVPPISFWRILKLNSSEWPYFVVGIFCAIVNGALQPAFSIIFSKVVGVFTRNT
                     DDETKRHDSNLFSLLFLILGVISFITFFLQGFTFGKAGEILTKRLRYMVFKSMLRQDV
                     SWFDNPKNTTGALTTRLANDAGQVKGATGARLAVITQNIANLGTGIIISLIYGWQLTL
                     LLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGSGKIATEAIENFRTVVSLTREQKFE
                     NMYAQSLQIPYRNALKKAHVFGITFSFTQAMMYFSYAACFRFGAYLVARELMTFENVL
                     LVFSAIVFGAMAVGQVSSFAPDYAKAKVSASHIIMIIEKVPSIDSYSTGGLKPNTLEG
                     NVKFNEVVFNYPTRPDIPVLQGLNLEVKKGQTLALVGSSGCGKSTVVQLLERFYDPMA
                     GTVFLDGKEVNQLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSQDEIERAA
                     KEANIHQFIESLPDKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEATSALD
                     TESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVIQNGKVKEHGTHQQLLAQKG
                     IYFSMVSVQAGAKR"
BASE COUNT     1275 a    895 c   1031 g   1103 t
ORIGIN      Map position 1q26.
```

```
Alignment Scores:
Pred. No.:              3.04e-223     Length:        4304
Score:                  3573.00       Matches:       679
Percent Similarity:     75.37%        Conservative:  248
Best Local Similarity:  55.20%        Mismatches:    259
Query Match:            59.43%        Indels:        44
DB:                     10            Gaps:          8

US-09-873-409-6 (1-1195) x CRUPGP1165 (1-4304)

Qy         1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
             |||:::::|||  ||||||::::::::::::|||    ||||||||||    |||
Db       258 ATGCTGGTGGGGACTCTCGCAGCTATTATCCATGGAGTTGCACTCCCGCTTATGATGCTG 317

Qy        21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsn------ 38
             |||   |||:::|||:::|||:::      |||          :::  |||||
Db       318 GTGTTTGGAGACATGACAGACAGCTTTGCAAGTGTAGGAAACATCCCAACAAATGCTACT 377

Qy        38 ------------------------------------------------------------ 38

Db       378 AATAATGCCACTCAAGTCAATGCTTCAGACATCTTTGGCAAACTGGAGGAAGAAATGACC 437

Qy        39 ThrTyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIle 58
             ||||||::::::               ||||||  |||||||||         ||||||
Db       438 ACGTACGCCTAT---------------TACTACACAGGGATTGGTGCTGGTGTGCTTATA 482

Qy        59 PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIle 78
                   ||||||||:::|||   |||   :::   |||  ||||||       :::|||
Db       483 GTTGCCTACATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATT 542

Qy        79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp 98
             |||::::::|||||||||:::::::::     |||:::|||||||||||||||    |||
Db       543 AGACAGAAGTTTTTCATGCTATAATGAATCAGGAGATAGGCTGGTTTGACGTGCACGAC 602

Qy        99 IleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly 117
             :::|||||||||||||||||||:::|||   |||:::   ||||||::::::||||||||
Db       603 GTTGGGGAGCTCAACACCCGGCTCACAGATGATGTCTCCAAAATTAACGAAGGAATTGGT 662

Qy       118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
             |||||||||   :::   ||||||   |||:::||||||      |||      :::|||
Db       663 GACAAAATTGGAATGTTCTTTCAGGCAATGGCAACATTTTTTGGTGGTTTTATAATAGGA 722

Qy       138 LeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
                :::||||||||||||||||||||   |||:::   ||||||::::::
Db       723 TTTACTCGAGGCTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTG 782

Qy       158 SerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr 177
             ||||||      ::::::::::::   |||   |||   |||||||||   ||||||
Db       783 TCAGCTGGTATTTGGGCAAAGATATTATCTTCATTTACTGATAAAGAACTTCAGGCATAT 842

Qy       178 SerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197
             :::||||||||||||||||||||||||||||||::::::||||||||||||||||||||||
Db       843 GCAAAAGCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCATTT 902

Qy       198 ArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 217
                |||:::|||||||||:::|||||||||      ||||||:::::::|||||
Db       903 GGAGGACAAAAGAAAGAACTTGAAAGGTACAACAACAATTTAGAAGAAGCCAAAAGACTT 962

Qy       218 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 237
             |||||||||:::   |||   :::   :::|||||||||   :::      :::
Db       963 GGGATAAAGAAAGCGATCACAGCCAACATTTCCATGGGCGCAGCTTTCCTTCTTATCTAT 1022

Qy       238 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 257
                :::|||  ||||||||||||||||||||||||||:::::::::     |||
Db      1023 GCATCATATGCTCTGGCCTTCTGGTATGGAACGTCCTTGGTCATCTCAAAAGAA------ 1076

Qy       258 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 277
             |||:::||||||   ||||||   |||||||||:::|||:::   :::   :::   |||
Db      1077 TATTCTATTGGACAAGTGCTCACTGTCTTCTTTGCTGTATTAATTGCGCCATTCAGTATT 1136

Qy       278 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 297
             |||   |||   |||:::   |||   ||||||   |||||||||||||||:::   |||
```

```
Db   1137 GGACAGGCATCTCCAAATATCGAAGCCTTTGCCAATGCAAGAGGAGCAGCCTATGAAATC 1196
Qy    298 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 317
Db   1197 TTCAACATAATTGATAATAAGCCCAGTATTGACAGCTTCTCAAAGAATGGGTACAAACCA 1256
Qy    318 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 337
Db   1257 GACAACATTAAAGGAAATTTGGAATTCAAAAATATTCACTTCAGTTACCCATCTCGAAAA 1316
Qy    338 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 357
Db   1317 GACGTTCAGATCTTGAAGGGCCTCAACCTGAAGGTGCAGAGCGGACAGACAGTGGCCCTG 1376
Qy    358 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 377
Db   1377 GTTGGCAACAGTGGCTGTGGGAAAAGCACCACTGTCCAGCTGCTGCAGAGGCTCTACGAC 1436
Qy    378 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 397
Db   1437 CCCACAGAGGGCGTGGTCAGTATCGACGGACAGGACATCAGGACCATCAATGTGAGGTAT 1496
Qy    398 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 417
Db   1497 CTGCGGGAAATCATTGGGGTCGTGAGTCAGGAACCTGTGTTGTTTGCCACCACTATCGCT 1556
Qy    418 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 437
Db   1557 GAAAACATTCGCTATGGCCGAGAAAATGTCACTATGGACGAGATCGAGAAAGCTGTCAAG 1616
Qy    438 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 457
Db   1617 GAAGCCAATGCCTATGACTTCATCATGAAACTGCCCCATAAATTTGACACTCTTGTTGGT 1676
Qy    458 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 477
Db   1677 GAGAGAGGAGCACAGCTGAGTGGAGGACAGAAACAGAGAATCGCCATTGCTCGTGCCCTG 1736
Qy    478 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 497
Db   1737 GTCCGCAACCCCAAGATCCTTTTGTTGGATGAGGCGACATCAGCCTTGGACACAGAAAGT 1796
Qy    498 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 517
Db   1797 GAAGCCGTGGTTCAGGCTGCTCTGGATAAGGCTAGAGAAGGCCGGACTACCATTGTGATA 1856
Qy    518 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 537
Db   1857 GCCCACCGCTTGTCTACAGTTCGAAATGCTGACATCATAGCTGGGTTTGATGGTGGTGTC 1916
Qy    538 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 557
Db   1917 ATTGTGGAGCAAGGAAATCATGAAGAGCTCATGAGAGAGAAGGGCATTTACTTCAAACTT 1976
Qy    558 ValMetSerGln-----------------------------AspIleLysLysAlaAsp 567
Db   1977 GTCATGACACAGACAGCAGGAAATGAAATTGAATTAGGAAATGAAGTTGGTGAGTCTAAA 2036
Qy    568 GluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu------Pro 585
Db   2037 AATGAAATTGATAATTTAGACATGTCTTCAAAAGATTCAGCATCCAGTCTAATCAGAAGG 2096
Qy    586 LeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr--- 604
Db   2097 AGATCAACTCGCAGAAGTATCCGTGGACCACATGACCAAGACAGGAAGCTTAGTACCAAA 2156
Qy    605 GlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLys 624
Db   2157 GAGGCCTTGGATGAAGATGTACCTCCAATTTCCTTTTGGCGTATCCTAAAGTTGAATTCA 2216
Qy    625 ProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHis 644
Db   2217 TCTGAATGGCCATATTTTGTGGTTGGTATATTCTGTGCCATAGTAAATGGAGCCTTGCAA 2276

Qy    645 ProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLys 663
Db   2277 CCAGCATTCTCAATAATATTTTCTAAGGTTGTAGGGGTTTTCACAAGAAATACTGATGAT 2336
Qy    664 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 683
Db   2337 GAAACCAAACGACATGATAGCAACTTGTTTTCCTTATTGTTTCTGATCCTTGGGGTCATT 2396
Qy    684 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 703
Db   2397 TCTTTTATTACATTTTTTCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACC 2456
Qy    704 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 723
Db   2457 AAGCGACTCCGATACATGGTTTTCAAATCCATGCTGAGACAGGATGTGAGCTGGTTTGAT 2516
Qy    724 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 743
Db   2517 AACCCTAAAAACACCACTGGAGCATTGACCACCAGGCTCGCCAACGATGCTGGTCAAGTT 2576
Qy    744 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 763
Db   2577 AAAGGGGCTACAGGAGCCAGACTTGCTGTCATTACCCAGAATATAGCAAATCTTGGGACA 2636
Qy    764 SerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAla 783
Db   2637 GGAATCATCATATCCCTAATCTATGGCTGGCAGTTGACACTTCTACTCTTAGCAATTGTC 2696
Qy    784 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 803
Db   2697 CCCATCATTGCAATAGCAGGAGTGGTTGAGATGAAAATGTTGTCTGGACAAGCACTAAAA 2756
Qy    804 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 823
Db   2757 GATAAGAAGGAGCTAGAAGGTTCTGGGAAGATTGCTACTGAAGCAATAGAGAACTTCCGC 2816
Qy    824 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 843
Db   2817 ACTGTCGTCTCTTTGACTCGGGAGCAGAAGTTTGAAAATATGTATGCCCAGAGCTTGCAG 2876
Qy    844 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 863
Db   2877 ATACCATACAGAAATGCTCTGAAGAAAGCACACGTCTTTGGGATCACATTCTCCTTCACC 2936
Qy    864 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 883
Db   2937 CAGGCCATGATGTATTTTTCGTATGCTGCTTGTTTCCGGTTTGGTGCTTACTTGGTGGCA 2996
Qy    884 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 903
Db   2997 CGGGAATTAATGACATTTGAAAATGTTCTATTAGTATTCTCAGCTATTGTCTTTGGTGCC 3056
Qy    904 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 923
Db   3057 ATGGCAGTTGGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTGTCGGCA 3116
Qy    924 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 943
Db   3117 TCGCATATCATCATGATCATTGAAAAAGTCCCTTCCATTGACAGCTACAGCACCGGAGGC 3176
Qy    944 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 963
Db   3177 CTGAAGCCTAATACATTGGAAGGAAATGTGAAATTTAATGAAGTCGTGTTCAACTACCCC 3236
Qy    964 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 983
Db   3237 ACCCGGCCTGACATCCCAGTGCTTCAGGGGCTGAACCTGGAGGTGAAGAAGGGCCAGACG 3296
Qy    984 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 1003
Db   3297 CTGGCCCTGGTGGGCAGCAGTGGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAGCGG 3356
```

```
Qy   1004 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 1023
             |||||||||:::  |||  |||     |||||   :::     :::|||||
Db   3357 TTCTACGACCCCATGGCCGGAACAGTGTTTCTAGATGGCAAAGAAGTAAACCAACTAAAT 3416

Qy   1024 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 1043
          |||||||||||||||:::   :::  ||||||  |||||||||:::||||||:::|||
Db   3417 GTCCAGTGGCTCCGAGCACACCTAGGCATTGTGTCCCAAGAGCCCATCCTGTTTGACTGC 3476

Qy   1044 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 1063
          ||||||||||||||||||||||||||||||||||||||||||||||    ||||||||
Db   3477 AGCATCGCTGAGAACATCGCTTATGGAGACAACAGCCGGGTCGTGTCCCAGGATGAGATC 3536

Qy   1064 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 1083
          :::   |||||      |||||||||||   ||||||||   ||||||:::||||||
Db   3537 GAGAGGGCGGCGAAGGAAGCCAACATCCACCAGTTCATTGAGTCACTGCCTGACAAATAC 3596

Qy   1084 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 1103
          ||||||:::||||||   ||||||   |||||||||||||||||||||||||:::|||
Db   3597 AACACCAGAGTAGGAGACAAAGGGACTCAGCTGTCGGGTGGGCAGAAACAACGCATCGCC 3656

Qy   1104 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 1123
          |||||||||||||||:::::::::|||   ||||||||||||||||||||||||||||
Db   3657 ATAGCACGCGCCCTCGTCAGACAGCCTCACATTTTGCTCCTGGATGAAGCAACATCAGCC 3716

Qy   1124 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 1143
          ||||||   :::|||||||||||||||||   |||||||||||||||||   ||||||
Db   3717 CTGGATACAGAGAGTGAAAAGGTTGTCCAGGAAGCTCTGGACAAAGCCAGAGAAGGCCGC 3776

Qy   1144 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 1163
          ||||||:::|||:::   ||||||||||||   |||||||||||||||||||||||||||
Db   3777 ACCTGCATTGTGATCGCTCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTG 3836

Qy   1164 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 1183
          :::   |||||||||:::||||||   ||||||||||||:::||||||      :::
Db   3837 ATTCAGAATGGCAAGGTCAAGGAGCATGGCACCCACCAGCAGCTGCTGGCACAGAAAGGC 3896

Qy   1184 IleTyrPheLysLeuValAsnAlaGlnSer 1193
          |||||||||   :::|||:::   |||:::
Db   3897 ATCTATTTCTCCATGGTCAGTGTGCAGGCT 3926
```

```
RESULT 8
AX105082
LOCUS       AX105082                4279 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 26 from Patent WO0123540.
ACCESSION   AX105082
VERSION     AX105082.1  GI:13921232
KEYWORDS    .
SOURCE      dog.
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 4279)
  AUTHORS   Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and
            Patten,C.J.
  TITLE     P-glycoproteins and uses thereof
  JOURNAL   Patent: WO 0123540-A 26 05-APR-2001;
            GENTEST CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..4279
                     /organism="Canis familiaris"
                     /db_xref="taxon:9615"
     CDS             17..3862
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC37723.1"
                     /db_xref="GI:13921233"
                     /translation="MDPEGGRKGSAEKNFWKMGKKSKKKEKKEKKPTVSTFAMFRYSN
                     WLDRLYMLVGTMAAIIHGAALPLMMLVFGNMTDSFANAGISRNKTFPVIINESITNNT
                     QHFINHLEEEMTTYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQILKIRKQFFHAIMRQ
                     EIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSIATFFTGFIVGFTRGWKLTL
                     VILAISPVLGLSAAIWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFGGQKKEL
                     ERYNKNLEEAKGIGIKKAITANISIGAAFLLIYASYALAFWYGTSLVLSSEYTIGQVL
                     TVFFSVLIGAFSIGQASPSIEAFANARGAAYEIFKIIDNKPSIDSYSKSGHKPDNIKG
                     NLEFKNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQRLYDPTD
                     GMVCIDGQDIRTINVRHLREITGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKE
                     ANAYDFIMKLPNKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTE
                     SEAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELMKEKGIY
                     FKLVTMQTRGNEIELENATGESKSESDALEMSPKDSGSSLIKRRSTRRSIHAPQGQDR
                     KLGTKEDLNENVPPVSFWRILKLNSTEWPYFVVGIFCAIINGGLQPAFSIIFSRIIGI
                     FTRDEDPETKRQNSNMFSVLFLVLGIISFITFFLQGFTFGKAGEILTKRLRYMVFRSM
                     LRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAVITQNIANLGTGIIISLIYG
                     WQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRTVVSLTR
                     EQKFEYMYAQSLQVPYRNSLRKAHIFGVSFSITQAMMYFSYAGCFRFGAYLVANEFMN
                     FQDVLLVFSAIVFGAMAVGQVSSFAPDYAKAKVSAAHVIMIIEKSPLIDSYSPHGLKP
                     NTLEGNVTFNEVVFNYPTRPDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERF
                     YDPLAGSVLIDGKEIKHLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSHEE
                     IVQAAKEANIHHFIETLPEKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEA
                     TSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGKVKEHGTHQQL
                     LAQKGIYFSMVSVQAGAKR"
BASE COUNT     1296 a    833 c   1009 g   1141 t
ORIGIN

Alignment Scores:
Pred. No.:              3.32e-222      Length:         4279
Score:                  3557.00        Matches:        683
Percent Similarity:     74.74%         Conservative:   237
Best Local Similarity:  55.48%         Mismatches:     269
Query Match:            59.17%         Indels:         42
DB:                     6              Gaps:           8

US-09-873-409-6 (1-1195) x AX105082 (1-4279)
```

```
Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          |||:::::|||   :::|||:::::::::::|||||   ||||||||||||   |||
Db    167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
          |||   |||   |||:::|||:::   :::
Db    227 GTTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy     35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                               :::   ||||||   |||
Db    287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy     45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                         ||||||   |||||||||         |||:::
Db    347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy     60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
             |||||||||:::|||   |||   :::   |||   ||||||      :::||||||
Db    407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy     80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
          ||||||||||||||:::::::::   |||:::|||||||||||||||     |||:::
Db    467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy    100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
          ||||||||||||||||||:::|||   |||:::   ||||||::::::|||||||||||
Db    527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy    119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
          ||||||   :::   ||||||:::::::::||||||         |||      ||||||
Db    587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy    139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
             :::|||||||||||||||||||||   |||:::   ||||||::::::      |||
Db    647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy    159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
          ||||||      ::::::::::::   |||   |||   |||||||||   ||||||:::
Db    707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy    179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
          |||||||||||||||||||||||||||::::::||||||||||||||||||||||
Db    767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826
```

```
Qy   199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
Db   827 GGACAAAAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAGGAATTGGG 886

Qy   219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238
Db   887 ATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCA 946

Qy   239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
Db   947 TCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TAT 1000

Qy   259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
Db  1001 ACTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGA 1060

Qy   279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
Db  1061 CAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTC 1120

Qy   299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
Db  1121 AAGATAATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGAT 1180

Qy   319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
Db  1181 AATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAA 1240

Qy   339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
Db  1241 GTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTT 1300

Qy   359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
Db  1301 GGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCC 1360

Qy   379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
Db  1361 ACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTT 1420

Qy   399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
Db  1421 CGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAA 1480

Qy   419 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
Db  1481 AACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAA 1540

Qy   439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
Db  1541 GCCAATGCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAG 1600

Qy   459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
Db  1601 AGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTT 1660

Qy   479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
Db  1661 CGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAA 1720

Qy   499 SerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 518
Db  1721 GCAGTGGTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCT 1780

Qy   519 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 538
Db  1781 CATCGTTTGTCTACAGTTCGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATT 1840

Qy   539 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 558
Db  1841 GTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTC 1900

Qy   559 MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 578
Db  1901 ACAATGCAG---ACAAGAGGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAA 1957

Qy   579 ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--- 597
Db  1958 AGTGAAAGTGATGCCTTGGAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAA 2014

Qy   598 ----------------------------------------AspLysAlaGluGluSer 603
Db  2015 AGAAGATCAACTCGCAGGAGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACA 2074

Qy   604 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 623
Db  2075 AAAGAGGACTTGAATGAGAATGTACCTCCAGTTTCCTTCTGGAGGATTCTGAAGCTGAAC 2134

Qy   624 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 643
Db  2135 TCAACTGAATGGCCTTATTTTGTGGTTGGTATATTTTGTGCTATTATAAACGGAGGCCTG 2194

Qy   644 HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAsp 662
Db  2195 CAACCAGCATTTTCAATAATATTTTCAAGGATTATAGGATCTTTACCCGAGATGAGGAT 2254

Qy   663 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 682
Db  2255 CCTGAAACAAAACGACAGAATAGTAACATGTTTTCTGTATTGTTTCTAGTCCTTGGAATT 2314

Qy   683 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 702
Db  2315 ATTTCTTTTATTACATTTTTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTC 2374

Qy   703 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 722
Db  2375 ACTAAGCGGCTTCGATACATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTT 2434

Qy   723 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 742
Db  2435 GATGACCCTAAAAACACCACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAA 2494

Qy   743 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 762
Db  2495 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGG 2554

Qy   763 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 782
Db  2555 ACAGGCATTATTATATCCTTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATT 2614

Qy   783 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 802
Db  2615 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 2674

Qy   803 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 822
Db  2675 AAAGATAAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTC 2734

Qy   823 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 842
Db  2735 CGAACTGTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTG 2794

Qy   843 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 862
Db  2795 CAAGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATC 2854

Qy   863 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 882
Db  2855 ACCCAGGCAATGATGTATTTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTG 2914

Qy   883 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 902
Db  2915 GCAAATGAGTTCATGAACTTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGT 2974

Qy   903 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 922
```

```
Db   2975 GCCATGGCAGTGGGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCA 3034

Qy    923 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 942
          |||||||||:::     ::::::||||||   |||   |||||||||   |||
Db   3035 GCAGCCCACGTCATCATGATCATTGAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCAC 3094

Qy    943 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 962
          |||   ||||||:::|||   |||||||||:::   |||   ||||||   |||   |||
Db   3095 GGCCTCAAGCCAAATACGTTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTAT 3154

Qy    963 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 982
          |||   |||   :::|||:::|||||||||||||         ::::::::|||:::
Db   3155 CCCACTCGACCAGACATCCCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAG 3214

Qy    983 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 1002
          |||:::|||   |||||||||||||||||||||||||||||||   ||||||||||||:::
Db   3215 ACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAG 3274

Qy   1003 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 1022
          |||   |||||||||:::   |||   ||||||   ||||||   :::   |||   |||
Db   3275 CGCTTCTATGACCCCTTGGCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTG 3334

Qy   1023 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 1042
          ||||||||||||||||||:::   :::   ||||||   |||||||||:::||||||:::
Db   3335 AATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGAC 3394

Qy   1043 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 1062
          ||||||||||||||||||||||||||||||||||||||||||||||||      :::|||
Db   3395 TGCAGCATTGCCGAGAACATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAG 3454

Qy   1063 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 1082
          |||   :::||||||      ||||||||||||   |||||||||   ||||||||||||
Db   3455 ATTGTGCAGGCAGCCAAGGAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAA 3514

Qy   1083 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1102
          ||||||||||:::||||||   ||||||   ||||||||||||||||||||||||:::
Db   3515 TACAACACCAGAGTAGGAGACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATT 3574

Qy   1103 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 1122
          ||||||||||||||||||:::::::::|||   ||||||||||||||||||||||||||||
Db   3575 GCCATAGCTCGCGCTCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCTACATCA 3634

Qy   1123 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1142
          |||||||||   :::||||||||||||||||||   ||||||||||||||||||   |||
Db   3635 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 3694

Qy   1143 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1162
          |||||||||:::|||:::   ||||||   ||||||||||||||||||||||||||
Db   3695 CGCACCTGCATTGTGATCGCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTG 3754

Qy   1163 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1182
          |||   |||||||||:::||||||   |||||||||||||:::||||||      :::
Db   3755 GTGTTTCAGAATGGCAAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGGCCCAGAAA 3814

Qy   1183 AspIleTyrPheLysLeuValAsnAlaGlnSer 1193
             |||||||||   :::|||:::   |||:::
Db   3815 GGCATCTATTTTTCCATGGTCAGTGTCCAGGCT 3847
```

RESULT 9
AX108656

```
LOCUS       AX108656                4195 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 3 from Patent WO0123565.
ACCESSION   AX108656
VERSION     AX108656.1  GI:13923888
KEYWORDS    .
SOURCE      crab-eating macaque.
  ORGANISM  Macaca fascicularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 4195)
  AUTHORS   Stocker,P.J., Steimel-Crespi,D.T. and Crespi,C.L.
  TITLE     P-glycoproteins from macaca fascicularis and uses thereof
  JOURNAL   Patent: WO 0123565-A 3 05-APR-2001;
            GENTEST CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..4195
                     /organism="Macaca fascicularis"
                     /db_xref="taxon:9541"
     CDS             100..3951
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC37765.1"
                     /db_xref="GI:13923889"
                     /translation="MDLEGDRNGGAEKKNFFKLNNKSKKDKKERKPTVSVFSMFRYSN
                     WLDKLYMVVGTLAAIIHGAGLPLMMLVFGDMTDTFANAGNLGDLGALLFNNTNSSNIT
                     DTVPVMNLEEDMTRYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQIHKIRKQFFHAIMR
                     QEIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSMATFFTGFIVGFTRGWKLT
                     LVILAISPVLGLSAAVWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFGGQKKE
                     LERYNKNLEEAKRIGIKKAITANISIGAAFLLIYASYALAFWYGTTLVLSKEYSIGQV
                     LTVFFSVLIGAFSVGQASPSIEAFANARGAAFEIFKIIDNKPSIDSYSKSGHKPDNIK
                     GNLEFRNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQRLYDPT
                     EGMVSVDGQDIRTINVRFLREIIGVVSQEPVLFATTIAENIRYGREDVTMDEIEKAVK
                     EANAYDFIMKLPQKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDT
                     ESEAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELMKEKGI
                     YFKLVTMQTAGNEIELENAADESKSEIDTLEMSSHDSGSSLIRKRSTRRSVRGSQGQD
                     RKLSTKEALDESIPPVSFWRIMKLNLTEWPYFVVGVFCAIINGGLQPAFAVIFSKIIG
                     IFTRNDDAETKRQNSNLFSLLFLVLGIVSFITFFLQGFTFGKAGEILTKRLRYMVFRS
                     MLRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAIITQNIANLGTGIIISLIY
                     GWQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRTVVSLT
                     QEQKFEHMYDQSLQVPYRNSLRKAHIFGITFSFTQAMMYFSYAGCFRFGAYLVAHSLM
                     SFEDVLLVFSAVVFGAMAVGQVSSFAPDYAKAKVSAAHIIMIIEKTPLIDSYSTEGLK
                     PNTLEGNVTFNEVVFNYPTRLDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLER
                     FYDPLAGKVLLDGKEIKQLNVQWLRAHLGIVSQEPILFDCSISENIAYGDNSRVVSQE
                     EIVRAAKEANIHAFIESLPNKYSTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDE
                     ATSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGRVKEHGTHQQ
                     LLAQKGIYFSMVSVQAGAKRQ"
BASE COUNT     1230 a    801 c    1039 g    1125 t
ORIGIN
```

Alignment Scores:

Pred. No.:	3.49e-222	Length:	4195
Score:	3556.50	Matches:	680
Percent Similarity:	75.16%	Conservative:	246
Best Local Similarity:	55.19%	Mismatches:	263
Query Match:	59.16%	Indels:	43
DB:	6	Gaps:	9

US-09-873-409-6 (1-1195) x AX108656 (1-4195)

```
Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          |||::::::|||   ||||||::::::::::::||||||   ||||||||||||   |||
Db    250 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGAGCTGGACTTCCTCTCATGATGCTG 309

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSer---------------------GlyCys 33
          |||   |||:::|||:::|||            :::                  |||
Db    310 GTGTTTGGAGACATGACGGATACCTTTGCAAATGCAGGAAATTTAGGAGATTTAGGAGCT 369

Qy     34 LeuValGlnThrAsnThrTyrSer------------------------------------ 41
          |||:::      ||||||   |||
Db    370 CTGTTGTTTAACAACACTAATAGCAGTAATATCACTGATACAGTGCCCGTCATGAATCTG 429

Qy     42 ---------PhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIle 58
                   |||         ||||||   |||||||||         |||:::
Db    430 GAGGAAGATATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTT 489

Qy     59 PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIle 78
             |||||||||:::|||   |||   :::   |||   ||||||      :::|||
Db    490 GCTGCTTACATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATT 549

Qy     79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp 98
          ||||||||||||||||||:::::::::   |||:::||||||||||||||||      |||
Db    550 AGAAAACAGTTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGAT 609

Qy     99 IleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly 117
          :::|||||||||||||||||||:::|||   |||:::   ||||||::::::|||||||||
```

```
Db    610 GTTGGGGAGCTTAACACCCGGCTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGT 669

Qy    118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
Db    670 GACAAAATTGGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGA 729

Qy    138 LeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
Db    730 TTTACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTG 789

Qy    158 SerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr 177
Db    790 TCAGCTGCAGTCTGGGCAAAGATACTGTCTTCATTTACTGATAAAGAACTCTTAGCTTAT 849

Qy    178 SerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197
Db    850 GCAAAAGCTGGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTT 909

Qy    198 ArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 217
Db    910 GGAGGACAAAAGAAAGAACTCGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATT 969

Qy    218 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 237
Db    970 GGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTTATCTAT 1029

Qy    238 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 257
Db   1030 GCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAAAGGAA------ 1083

Qy    258 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 277
Db   1084 TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTT 1143

Qy    278 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 297
Db   1144 GGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTTTGAAATC 1203

Qy    298 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 317
Db   1204 TTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCA 1263

Qy    318 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 337
Db   1264 GATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAA 1323

Qy    338 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 357
Db   1324 GAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTG 1383

Qy    358 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 377
Db   1384 GTTGGAAACAGCGGCTGTGGGAAGAGCACAACGGTCCAGCTGATGCAGAGGCTTTATGAC 1443

Qy    378 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 397
Db   1444 CCCACAGAGGGCATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAACGTAAGGTTT 1503

Qy    398 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 417
Db   1504 CTACGGGAAATCATCGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCT 1563

Qy    418 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 437
Db   1564 GAAAACATTCGCTATGGTCGTGAAGATGTCACCATGGATGAGATTGAGAAAGCTGTCAAG 1623

Qy    438 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 457
Db   1624 GAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCAGAAATTTGACACCCTGGTTGGA 1683

Qy    458 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 477
Db   1684 GAGAGAGGGGCCCAGCTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTG 1743

Qy    478 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 497
Db   1744 GTTCGCAACCCCAAGATCCTCCTGCTGGACGAGGCCACGTCAGCCTTGGACACAGAAAGT 1803

Qy    498 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 517
Db   1804 GAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATA 1863

Qy    518 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 537
Db   1864 GCTCATCGTTTGTCTACGGTTCGTAATGCCGACGTCATCGCTGGTTTCGATGATGGAGTC 1923

Qy    538 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 557
Db   1924 ATTGTGGAGAAAGGAAATCATGATGAGCTCATGAAAGAGAAAGGCATTTACTTCAAACTT 1983

Qy    558 ValMetSerGlnAspIle--------------------LysLysAlaAspGlu------ 568
Db   1984 GTCACAATGCAGACAGCAGGAAATGAAATTGAATTAGAAAATGCAGCTGATGAATCCAAA 2043

Qy    569 ---GlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu--------- 584
Db   2044 AGTGAAATTGATACCTTGGAAATGTCTTCACATGATTCAGGATCCAGTCTAATAAGAAAA 2103

Qy    585 ------ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGlu 602
Db   2104 AGATCCACTCGTAGGAGTGTCCGTGGATCACAAGGCCAA------GACAGAAAGCTTAGT 2157

Qy    603 SerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeu 622
Db   2158 ACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTA 2217

Qy    623 AsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThr 642
Db   2218 AATTTAACTGAGTGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGT 2277

Qy    643 ValHisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsn 661
Db   2278 CTGCAACCAGCATTTGCAGTAATATTTTCAAAGATTATAGGGATTTTTACAAGAAATGAT 2337

Qy    662 AspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGly 681
Db   2338 GATGCCGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGTCCTTGGA 2397

Qy    682 ValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIle 701
Db   2398 ATTGTTTCTTTTATTACATTTTTCCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATC 2457

Qy    702 LeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrp 721
Db   2458 CTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGCTGG 2517

Qy    722 PheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAla 741
Db   2518 TTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCT 2577

Qy    742 GlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMet 761
Db   2578 CAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTATAATTACCCAGAATATAGCAAATCTT 2637

Qy    762 GlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSer 781
Db   2638 GGGACAGGAATAATTATATCCTTAATCTATGGTTGGCAACTGACACTGTTACTCTTAGCA 2697

Qy    782 IleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAla 801
Db   2698 ATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCA 2757

Qy    802 AsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsn 821
Db   2758 CTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAAC 2817
```

```
Qy    822 IleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMet 841
               ||||||:::||||||||||||:::|||:::    |||||    |||||::::::
Db   2818 TTCCGAACTGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGATCAGAGT 2877

Qy    842 LeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAla 861
          |||||        :::||||||:::   :::||||||   |||   |||       ::::::
Db   2878 TTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATCACGTTTTCC 2937

Qy    862 PheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeu 881
          |||:::   |||   :::||||||:::||||||         ||||||||||||||||||
Db   2938 TTCACGCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTG 2997

Qy    882 IleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyr 901
          :::               |||:::   |||   :::   :::||||||:::|||:::   :::
Db   2998 GTGGCACATAGTCTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTT 3057

Qy    902 GlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSer 921
          ||||||||||||:::|||:::            ||||||:::|||:::|||||||
Db   3058 GGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTA 3117

Qy    922 GlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGln 941
             |||||||||:::      ::::::||||||   |||   |||||||||   |||
Db   3118 TCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACA 3177

Qy    942 GluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePhe 961
             |||||   ||||||:::|||   |||||||||:::   |||   ||||||   |||
Db   3178 GAAGGCCTAAAGCCGAACACATTGGAAGGAAATGTCACATTTAATGAAGTTGTATTCAAC 3237

Qy    962 TyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGly 981
          ||||||   |||   |||:::   :::|||:::||||||||||||   :::::::::|||
Db   3238 TATCCCACCCGACTGGACATCCCAGTGCTTCAGGGGCTGAGCCTGGAAGTGAAGAAGGGC 3297

Qy    982 LysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeu 1001
          :::|||:::|||   ||||||||||||||||||||||||||||   ||||||||||||
Db   3298 CAGACGCTGGCCCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACGGTGGTCCAGCTCCTG 3357

Qy   1002 GlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGlu 1021
          :::|||   |||||||||   |||:::||||||   ||||||      :::   |||:::
Db   3358 GAGCGGTTCTATGACCCCTTGGCGGGGAAAGTGCTGCTTGACGGCAAAGAAATAAAGCAA 3417

Qy   1022 LeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPhe 1041
          |||||||||||||||||||||:::      :::   ||||||   |||||||||:::||||||
Db   3418 CTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTT 3477

Qy   1042 AsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAsp 1061
          :::||||||||:::||||||||||||||||||||||||||||||||||||         :::
Db   3478 GACTGCAGCATTAGTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAA 3537

Qy   1062 GluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGlu 1081
          ||||||      ||||||      ||||||||||||:::|||||||||   ||||||
Db   3538 GAGATCGTGAGGGCAGCCAAGGAGGCCAATATACACGCCTTCATCGAGTCACTGCCTAAT 3597

Qy   1082 LysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArg 1101
          ||||||:::|||:::||||||   ||||||   ||||||||||||||||||||||||
Db   3598 AAATATAGCACCAGAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGC 3657

Qy   1102 LeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThr 1121
          :::||||||||||||||||||:::::::::|||   ||||||||||||||||||||||||
Db   3658 ATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACA 3717

Qy   1122 SerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThr 1141
          ||||||||||||      :::||||||||||||||||||   ||||||||||||||||
Db   3718 TCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAA 3777

Qy   1142 GlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIle 1161
          ||||||||||||:::|||:::   ||||||||||||   ||||||||||||||||||||||
Db   3778 GGCCGTACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATA 3837

Qy   1162 ValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsn 1181
          ||||||      ||||||::::::||||||   ||||||||||||:::||||||
Db   3838 GTGGTGTTTCAGAATGGCAGAGTCAAGGAGCACGGCACACATCAGCAGCTGCTGGCACAG 3897

Qy   1182 ArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1193
              :::   ||||||||   :::|||:::   |||:::
Db   3898 AAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3933
```

```
RESULT 10
AF257746
LOCUS       AF257746                4927 bp    mRNA    linear   ROD 09-MAY-2000
DEFINITION  Rattus norvegicus multidrug resistance protein 1a (Pgy1) mRNA,
            complete cds.
ACCESSION   AF257746
VERSION     AF257746.1  GI:7739772
KEYWORDS    .
SOURCE      Rattus norvegicus.
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
REFERENCE   1  (bases 1 to 4927)
  AUTHORS   Hooiveld,G.J.E.J., Wilms,J.W.J., Hagenbuch,B., Jansen,P.L.M.,
            Meijer,D.K.F. and Muller,M.
  TITLE     Cloning and functional characterization of the rat multidrug
            resistance protein Mdr1a
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4927)
  AUTHORS   Hooiveld,G.J.E.J., Wilms,J.W.J., Hagenbuch,B., Jansen,P.L.M.,
            Meijer,D.K.F. and Muller,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-APR-2000) Div. Gastroenterology & Hepatology,
            University Hospital Groningen, Hanzeplein 1, Groningen NL-9713 GZ,
            The Netherlands
FEATURES             Location/Qualifiers
     source          1..4927
                     /organism="Rattus norvegicus"
                     /strain="Wistar"
                     /db_xref="taxon:10116"
                     /sex="male"
                     /cell_type="hepatocyte"
                     /tissue_type="liver"
                     /dev_stage="adult"
     gene            1..4927
                     /gene="Pgy1"
     CDS             86..3904
                     /gene="Pgy1"
                     /note="ATP-binding cassette transport protein; Abcb1a"
                     /codon_start=1
                     /product="multidrug resistance protein 1a"
                     /protein_id="AAF69007.1"
                     /db_xref="GI:7739773"
                     /translation="MELEEDLNGRADKNFSKMGKKSKKEKKEKKPAVSVLTMFRYAGW
                     LDRFYMLLGTLAAIIHGIALPLMMLVFGDMTDSFANVGNNRSMSFYNATDIYAKLEDE
                     MTTYAYYYTGIGAGVLIVAYIQVSLWCLAAGRQIHKIRQKFFHAIMNQEIGWFDVHDV
                     GELNTRLTDDVSKINEGIGDKIGMFFQAMATFFGGFIIGFTRGWKLTLVILAISPVLG
                     LSAGIWAKILSSFTDKELQAYAKAGAVAEEVLAAIRTVIAFGGQKKELERYNNNLEEA
                     KRLGIKKAITANISMGAAFLLIYASYALAFWYGTSLVISKEYTIGQVLTVFFSVLIGA
                     FSVGQASPNIEAFANARGAAYEVFSIIDNKPSIDSFSKSGHKPDNIQGNLEFKNIHFS
                     YPSRKDVQILKGLNLKVKSGQTVALVGNSGCGKSTTVQLLQRLYDPIEGEVSIDGQDI
                     RTINVRYLREIIGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKEANAYDFIMKL
                     PHKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTESEAVVQAALD
                     KAREGRTTIVIAHRLSTVRNADVIAGFDGGVIVEQGNHDELMREKGIYFKLVMTQTAG
                     NEIELGNEACESKDGIDNVDMSSKDSGSSLIRRRSTRKSIRGPHDQDGELSTKEALDD
                     DVPPASFWRILKLNSTEWPYFVVGVFCAIINGGLQPAFSIIFSKVVGVFTKNDTPEIQ
                     RQNSNLFSLLFLILGIISFITFFLQGFTFGKAGEILTKRLRYMVFKSMLRQDISWFDD
                     PKNTTGALTTRLANDAAQVKGATGSRLAVITQNIANLGTGIIISLIYGWQLTLLLLAI
                     VPIIAIAGVVEMKMLSGQALKDKKELEGSGKIATEAIENFRTVVSLTREQKFETMYAQ
                     SLQIPYRNALKKAHVFGITFSFTQAMMYFSYAACFRFGAYLVARELMTFENVLLVFSA
                     IVFGAMAVGQVSSFAPDYAKAKVSASHIIRIIEKIPEIDSYSTEGLKPNMLEGNVKFN
                     GVMFNYPTRPNIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERFYDPMAGTVFL
                     DGKEIKQLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSHEEIVKAAKEANI
                     HQFIDSLPEKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEATSALDTESEK
                     VVQEALDKAREGRTCIVIAHRLSTIQNADLIVVIQNGQVKEHGTHQQLLAQKGIYFSM
                     VSVQAGAKRS"
BASE COUNT     1474 a   1037 c   1190 g   1226 t
ORIGIN
```

```
Alignment Scores:
Pred. No.:              4.98e-222     Length:        4927
Score:                  3555.50       Matches:       682
Percent Similarity:     75.78%        Conservative:  244
Best Local Similarity:  55.81%        Mismatches:    263
Query Match:            59.14%        Indels:        33
DB:                     10            Gaps:          9

US-09-873-409-6 (1-1195) x AF257746 (1-4927)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    233 ATGCTGCTGGGAACTCTGGCGGCCATTATCCATGGAATTGCGCTCCCACTTATGATGCTG 292

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db    293 GTCTTTGGAGACATGACAGATAGCTTTGCAAAT------GTAGGAAACAACCGTAGTATG 346

Qy     41 SerPhePheArgLeuThr------------------------------------------ 46
Db    347 AGTTTCTACAATGCTACAGACATATATGCCAAGCTGGAGGACGAAATGACCACGTACGCC 406

Qy     47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
Db    407 TACTATTACACGGGCATTGGTGCCGGTGTGCTCATCGTTGCCTACATCCAGGTTTCACTT 466

Qy     67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
Db    467 TGGTGCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAAGTTTTTCCATGCCATC 526

Qy     87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
Db    527 ATGAATCAGGAGATAGGCTGGTTTGACGTGCATGACGTTGGGGAGCTCAACACCCGGCTC 586

Qy    107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
Db    587 ACAGATGACGTCTCCAAAATTAATGAAGGAATTGGTGACAAAATTGGAATGTTCTTTCAG 646

Qy    126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
Db    647 GCAATGGCAACATTTTTTGGTGGTTTTATAATAGGATTTACTCGCGGCTGGAAGCTAACT 706

Qy    146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
Db    707 CTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTGTCAGCTGGTATTTGGGCAAAGATA 766

Qy    166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
Db    767 TTGTCTTCATTTACTGATAAGGAACTCCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAA 826

Qy    186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
Db    827 GAAGTCTTAGCAGCCATCAGAACTGTGATTGCCTTTGGAGGACAAAAGAAGGAACTTGAA 886

Qy    206 ArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSer 225
Db    887 AGGTACAATAACAATTTGGAAGAAGCTAAAAGGCTTGGGATAAAGAAAGCTATCACGGCC 946

Qy    226 LysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrp 245
Db    947 AACATTTCCATGGGTGCAGCTTTTCTGCTTATCTATGCATCATATGCTCTGGCATTCTGG 1006

Qy    246 TyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAla 265
Db   1007 TATGGGACTTCCTTGGTCATCTCAAAAGAA------TACACTATTGGACAAGTGCTCACT 1060

Qy    266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285
Db   1061 GTCTTTTTTTCTGTATTAATTGGAGCATTCAGTGTTGGGCAGGCATCTCCAAATATTGAA 1120

Qy    286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305
Db   1121 GCCTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTCAGTATAATTGATAATAAGCCC 1180

Qy    306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
Db   1181 AGTATAGACAGCTTCTCAAAGAGTGGGCACAAACCCGACAACATACAAGGAAATTTGGAA 1240

Qy    326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
Db   1241 TTCAAAAATATTCACTTCAGTTACCCGTCTCGAAAAGACGTTCAGATCTTGAAGGGCCTC 1300

Qy    346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
Db   1301 AACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTGGTTGGCAACAGTGGCTGTGGGAAA 1360

Qy    366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
Db   1361 AGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCATAGAGGGCGAGGTCAGTATC 1420

Qy    386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
Db   1421 GACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTGGTG 1480

Qy    406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
Db   1481 AGTCAGGAACCCGTGCTGTTTGCCACCACAATTGCCGAAAACATTCGCTATGGCCGAGAA 1540

Qy    426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
Db   1541 AACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAATGCCTATGATTTCATC 1600

Qy    446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465
Db   1601 ATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGG 1660

Qy    466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
Db   1661 GGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAACCCCAAGATCCTTTTG 1720

Qy    486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeu 505
Db   1721 TTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTGGTTCAGGCCGCTCTG 1780

Qy    506 GluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArg 525
Db   1781 GATAAGGCTAGAGAAGGCCGGACCACCATTGTGATAGCTCACCGCTTGTCTACAGTTCGC 1840

Qy    526 SerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAla 545
Db   1841 AATGCTGACGTCATTGCTGGTTTTGATGGTGGTGTCATTGTGGAGCAAGGAAATCATGAT 1900

Qy    546 GluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln------------ 561
Db   1901 GAGCTCATGAGAGAGAAAGGAATTTACTTCAAACTTGTCATGACTCAGACAGCAGGAAAT 1960

Qy    562 AspIleLysLysAlaAspGluGlnMetGluSer------------------MetThrTyr 575
Db   1961 GAAATTGAATTAGGAAATGAAGCTTGTGAATCTAAAGATGGAATTGATAATGTGGACATG 2020

Qy    576 SerThrGluArgLysThrAsnSerLeu------ProLeuHisSerValLysSerIleLys 593
Db   2021 TCTTCAAAAGATTCAGGATCCAGTCTAATAAGAAGAAGATCAACTCGCAAAAGCATCCGT 2080

Qy    594 SerAspPheIleAspLysAlaGluGluSerThr---GlnSerLysGluIleSerLeuPro 612
Db   2081 GGGCCACATGATCAAGACGGGGAACTTAGCACCAAAGAGGCTCTGGATGACGACGTACCT 2140

Qy    613 GluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrpProPheValValLeu 632
Db   2141 CCAGCTTCCTTTTGGCGGATCCTGAAGTTGAATTCAACTGAATGGCCTTATTTTGTGGTT 2200

Qy    633 GlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPheSerIleIlePheAla 652
Db   2201 GGTGTATTTTGTGCCATAATAAATGGAGGCTTGCAACCAGCATTCTCCATAATATTTTCA 2260
```

```
Qy     653 LysIleIleThrMetPheGlyAsnAsnAspLysThrThrLeu---LysHisAspAlaGlu 671
           |||:::::  :::|||    ||||||           :::    :::   ::::::
Db    2261 AAGGTTGTAGGGGTTTTTACAAAAAATGACACCCCTGAAATCCAGCGGCAGAACAGCAAC 2320

Qy     672 IleTyrSerMetIlePheValIleLeuGlyValIleCysPheValSerTyrPheMetGln 691
           ::::::|||::::::|||:::||||||||:::|||   |||:::::::::|||:::|||
Db    2321 TTGTTTTCTTTATTGTTCCTGATCCTTGGGATCATCTCTTTCATTACGTTTTTCCTTCAG 2380

Qy     692 GlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeuArgHisLeuAlaPhe 711
           |||       :::|||:::||||||||||||||||||   ||||||||::::::   |||
Db    2381 GGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGACTCCGATACATGGTCTTC 2440

Qy     712 LysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGly 731
           |||:::||||||   |||||||||:::||||||||:::   :::|||:::||||||
Db    2441 AAATCCATGCTGAGACAGGACATAAGCTGGTTTGATGACCCTAAAAACACCACAGGAGCG 2500

Qy     732 LeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIle 751
           |||||||||   ||||||   |||   ||||||::::::||||||||||||||||||:::
Db    2501 CTGACCACCAGGCTTGCCAATGACGCTGCTCAAGTGAAAGGGGCTACAGGGTCTAGGCTT 2560

Qy     752 GlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyr 771
              |||:::|||||||||      |||:::|||      :::|||||||||   ||||||
Db    2561 GCTGTTATTACCCAGAACATAGCAAATCTTGGGACAGGCATCATCATATCCCTGATCTAC 2620

Qy     772 GlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeuAlaValThrGlyMet 791
           ||||||::::::|||   |||:::|||:::|||   |||::::::|||:::   |||:::
Db    2621 GGCTGGCAATTGACACTTTTACTCCTAGCAATTGTTCCCATCATTGCTATAGCAGGAGTG 2680

Qy     792 IleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAla 811
           :::|||         ::::::|||   |||   ||||||||:::||||||:::   :::
Db    2681 GTTGAAATGAAAATGTTGTCTGGACAAGCGCTGAAAGATAAGAAGGAACTAGAAGGTTCT 2740

Qy     812 GlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleValSerLeuThrArgGlu 831
           ||||||||||||||||||||:::||||||   ||||||:::||||||||||||||||||
Db    2741 GGGAAGATCGCTACAGAAGCAATTGAAAACTTTCGCACTGTCGTCTCTTTGACTCGGGAG 2800

Qy     832 LysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLys 851
           :::   ||||||   ||||||   :::   ||||||   :::||||||         |||
Db    2801 CAGAAGTTTGAAACTATGTATGCCCAGAGCTTGCAGATACCATACAGAAATGCTTTGAAG 2860

Qy     852 LysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyr 871
           ||||||   :::   |||      ::::::|||:::   |||   :::||||||:::|||
Db    2861 AAAGCGCACGTCTTTGGGATCACTTTCTCCTTCACCCAGGCCATGATGTATTTCTCCTAT 2920

Qy     872 AlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGly 891
           ||||||   |||||||||||||||||||||:::               ||||||   |||
Db    2921 GCTGCTTGTTTCCGGTTTGGTGCCTACTTGGTGGCACGAGAACTCATGACATTTGAAAAT 2980

Qy     892 MetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuVal 911
           :::   :::||||||:::||||||   :::||||||||||||:::|||:::
Db    2981 GTTCTGTTAGTATTCTCAGCTATTGTCTTTGGTGCCATGGCAGTGGGGCAGGTCAGTTCA 3040

Qy     912 LeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGlu 931
           ||||||:::|||:::|||||||||      |||:::         ::::::|||
Db    3041 TTCGCTCCTGACTACGCGAAAGCCAAAGTCTCGGCATCCCACATCATCAGGATCATTGAG 3100

Qy     932 LysLysProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGly 951
           |||   |||   |||||||||   |||   ||||||   ||||||:::      ||||||
Db    3101 AAAATCCCTGAGATTGACAGCTACAGCACGGAGGGCTTGAAGCCTAATATGTTGGAAGGA 3160

Qy     952 AsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeu 971
           |||::::::|||      |||   |||   ||||||   ||||||::::::   :::|||
Db    3161 AATGTGAAATTTAATGGAGTCATGTTCAACTATCCCACCCGACCCAACATCCCAGTGCTT 3220

Qy     972 ArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGly 991
           :::||||||||||||   :::::::::|||:::|||:::|||   |||||||||||||||
Db    3221 CAGGGGCTGAGCCTAGAGGTGAAGAAAGGGCAGACGCTGGCCCTCGTGGGCAGCAGTGGC 3280

Qy     992 CysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGln 1011
           ||||||||||||||||   ||||||||||||:::|||   |||||||||:::   |||
Db    3281 TGCGGGAAGAGTACAGTGGTCCAGCTGCTTGAGCGCTTCTATGACCCCATGGCCGGAACA 3340

Qy    1012 ValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIle 1031
           |||      ||||||   :::   |||:::||||||||||||||||||:::   :::
Db    3341 GTGTTTCTAGATGGCAAAGAAATAAAGCAACTCAATGTCCAGTGGCTCCGCGCCCACCTG 3400

Qy    1032 AlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyr 1051
              ||||||   |||||||||:::||||||:::||||||||||||||||||||||||
Db    3401 GGCATTGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATCGCCGAGAACATTGCCTAC 3460

Qy    1052 GlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsn 1071
           ||||||||||||||||||         :::||||||   :::||||||      ||||||
Db    3461 GGAGACAACAGCCGTGTCGTGTCTCATGAGGAGATCGTGAAGGCAGCCAAGGAGGCCAAC 3520

Qy    1072 IleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGly 1091
           ||||||   ||||||:::   ||||||||||||||||||:::||||||   ||||||
Db    3521 ATCCACCAGTTCATCGACTCACTGCCTGAGAAATACAACACCAGAGTGGGAGACAAAGGG 3580

Qy    1092 AlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLys 1111
              |||||||||||||||||||||||:::||||||||||||||||||:::::::::
Db    3581 ACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATCGCCATCGCGCGCGCCCTCGTCAGACAG 3640

Qy    1112 ProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysVal 1131
           |||   |||||||||||||||||||||||||||||||||   :::||||||||||||
Db    3641 CCTCACATCTTACTTCTGGATGAAGCGACATCAGCTCTGGATACGGAGAGTGAAAAGGTC 3700

Qy    1132 ValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArg 1151
           ||||||   ||||||||||||||||||   ||||||:::|||:::   ||||||
Db    3701 GTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACCTGCATTGTGATCGCGCACCGC 3760

Qy    1152 LeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGlu 1171
           ||||||   ||||||||||||||||||||||||:::   |||||||::::::||||||
Db    3761 CTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATTCAGAACGGCCAGGTCAAGGAG 3820

Qy    1172 GlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAla 1191
              ||||||||||||:::||||||      :::   ||||||||   :::|||:::
Db    3821 CACGGCACCCACCAGCAGCTGCTGGCCCAGAAAGGCATCTATTTCTCGATGGTCAGTGTG 3880

Qy    1192 GlnSer 1193
           |||:::
Db    3881 CAGGCT 3886
```

RESULT 11
AX322787

```
LOCUS       AX322787                3860 bp    DNA     linear   PAT 07-JAN-2002
DEFINITION  Sequence 1 from Patent WO0192877.
ACCESSION   AX322787
VERSION     AX322787.1  GI:18093766
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Sorrentino,B. and Schuetz,J.
  TITLE     Method of identifying and/or isolating stem cells
  JOURNAL   Patent: WO 0192877-A 1 06-DEC-2001;
            ST. JUDE CHILDREN'S RESEARCH HOSPITAL (US)
FEATURES             Location/Qualifiers
     source          1..3860
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1135 a    746 c    957 g   1022 t
ORIGIN

Alignment Scores:
Pred. No.:              4.56e-222       Length:         3860
Score:                  3554.00         Matches:        682
Percent Similarity:     75.63%          Conservative:   246
Best Local Similarity:  55.58%          Mismatches:     263
Query Match:            59.12%          Indels:         36
DB:                     6               Gaps:           9

US-09-873-409-6 (1-1195) x AX322787 (1-3860)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
```

```
Db    151 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210

Qy     21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db    211 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 270

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db    271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    331 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 390

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    451 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 510

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    511 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGGTATTGGTGACAAAATT 570

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 630

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    631 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    691 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db    811 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 870

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db    871 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 930

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db    931 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 984

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db    985 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1044

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db   1045 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db   1105 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1164

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db   1165 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1224

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   1225 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1284

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   1285 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1344

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   1345 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1404

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   1405 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1464

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1465 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1524

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1525 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1584

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   1585 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1644

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   1645 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1704

Qy    501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db   1705 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 1764

Qy    521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db   1765 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 1824

Qy    541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db   1825 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 1884

Qy    561 Gln------------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
Db   1885 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 1944

Qy    575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 589
Db   1945 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2004

Qy    590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
Db   2005 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2064

Qy    608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627
Db   2065 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2124

Qy    628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
Db   2125 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2184

Qy    648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
Db   2185 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2244

Qy    667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
Db   2245 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2304

Qy    687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
Db   2305 ACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2364
```

```
Qy    707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
          |||::::::  |||::::::|||||  |||||::::::||||||||:::   :::
Db   2365 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2424

Qy    727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
          |||:::||||||  |||||||||  ||||||  |||  ||||||::::::||||||
Db   2425 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2484

Qy    747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
             ||||||||:::  |||:::|||||||||     |||:::|||      :::|||
Db   2485 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2544

Qy    767 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
          |||||||||||||||||||||||:::::::|||  |||:::|||:::|||  |||::::::
Db   2545 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 2604

Qy    787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
          |||:::  |||::::::|||         ::::::|||  |||  ||||||||:::
Db   2605 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 2664

Qy    807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
          ||||||:::  ||||||||||||||||||||||||:::||||||  ||||||:::|||
Db   2665 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 2724

Qy    827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
          |||||||||:::|||:::  ||||||  ||||||   :::  ||||||      :::
Db   2725 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2784

Qy    847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
          ||||||:::  :::||||||  |||  |||    ::::::|||:::  |||
Db   2785 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 2844

Qy    867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
          :::||||||:::||||||       ||||||||||||||||||:::
Db   2845 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 2904

Qy    887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
          |||:::  |||   :::   :::||||||:::|||:::   :::||||||||||||:::
Db   2905 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 2964

Qy    907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
          |||:::             ||||||:::|||:::||||||||     |||||||||:::
Db   2965 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3024

Qy    927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
             ::::::||||||  |||  |||||||||  |||  |||              |||
Db   3025 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3084

Qy    947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
          :::|||     |||||||||:::  |||  |||  |||  ||||||  |||  ||||||
Db   3085 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3144

Qy    967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
          |||:::  :::|||:::||||||||||||   :::::::::|||:::|||:::|||
Db   3145 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3204

Qy    987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
          |||||||||||||||||||||||||||  ||||||||||||:::|||   ||||||
Db   3205 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3264

Qy   1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
          |||:::  |||:::||||||  ||||||    :::  |||  |||||||||||||||
Db   3265 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3324

Qy   1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
          ||||||:::   :::  ||||||  |||||||||:::||||||:::|||||||||||||
Db   3325 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3384

Qy   1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
          ||||||||||||||||||||||||||||||||||||        :::||||||     |||
Db   3385 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3444

Qy   1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
          |||        |||||||||||:::||||||||      ||||||    |||||:::|||:::
Db   3445 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3504

Qy   1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
          ||||||   ||||||  ||||||||||||||||||||||||||:::||||||||||||
Db   3505 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3564

Qy   1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
          ||||||:::::::::|||  ||||||||||||||||||||||||||||||||||||
Db   3565 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 3624

Qy   1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
          :::||||||||||||||||||   ||||||||||||||||||   ||||||||||||:::
Db   3625 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 3684

Qy   1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
          |||:::   |||||||||||  |||||||||||||||||||||||||||||   |||
Db   3685 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 3744

Qy   1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
          |||::::::||||||  |||||||||||:::||||||      :::  |||||||||
Db   3745 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 3804

Qy   1187 LysLeuValAsnAlaGlnSer 1193
             :::|||:::   |||:::
Db   3805 TCAATGGTCAGTGTCCAGGCT 3825
```

```
RESULT 12
AX105078
LOCUS       AX105078                 4279 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 22 from Patent WO0123540.
ACCESSION   AX105078
VERSION     AX105078.1  GI:13921228
KEYWORDS    .
SOURCE      dog.
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 4279)
  AUTHORS   Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and
            Patten,C.J.
  TITLE     P-glycoproteins and uses thereof
  JOURNAL   Patent: WO 0123540-A 22 05-APR-2001;
            GENTEST CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..4279
                     /organism="Canis familiaris"
                     /db_xref="taxon:9615"
     CDS             17..3862
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC37721.1"
                     /db_xref="GI:13921229"
                     /translation="MDPEGGRKGSAEKNFWKMGKKSKKNEKKEKKPTVSTFAMFRYSN
                     WLDRLYMLVGTMAAIIHGAALPLMMLVFGNMTDSFANAGISRNKTFPVIINESITNNT
                     QHFINHLEEEMTTYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQILKIRKQFFHAIMRQ
                     EIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSIATFFTGFIVGFTRGWKLTL
                     VILAISPVLGLSAAIWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFGGQKKEL
                     ERYNKNLEEAKGIGIKKAITANISIGAAFLLIYASYALAFWYGTSLVLSSEYSIGQVL
                     TVFFSVLIGAFSIGQASPSIEAFANARGAAYEIFKIIDNKPSIDSYSKSGHKPDNIKG
                     NLEFKNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQRLYDPTD
                     GMVCIDGQDIRTINVRHLREITGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKE
                     ANAYDFIMKLPNKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTE
                     SEAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELMKEKGIY
                     FKLVTMQTRGNEIELENATGESKSESDALEMSPKDSGSSLIKRRSTRRSIHAPQGQDR
                     KLGTKEDLNENVPPVSFWRILKLNSTEWPYFVVGIFCAIINGGLQPAFSIIFSRIIGI
                     FTRDEDPETKRQNSNMFSVLFLVLGIISFITFFLQGFTFGKAGEILTKRLRYMVFRSM
                     LRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAVITQNIANLGTGIIISLIYG
                     WQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRTVVSLTR
                     EQKFEYMYAQSLQVPYRNSLRKAHIFGVSFSITQAMMYFSYAGCFRFGAYLVANEFMN
                     FQDVLLVFSAIVFGAMAVGQVSSFAPDYAKAKVSAAHVIMIIEKSPLIDSYSPHGLKP
                     NTLEGNVTFNEVVFNYPTRPDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERF
```

```
                     YDPLAGSVLIDGKEIKHLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSHEE
                     IMQAAKEANIHHFIETLPEKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEA
                     TSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGKVKEHGTHQQL
                     LAQKGIYFSMVSVQAGAKR"
BASE COUNT     1295 a    833 c   1008 g   1143 t
ORIGIN

Alignment Scores:
Pred. No.:              5.21e-222      Length:        4279
Score:                  3554.00        Matches:       682
Percent Similarity:     74.74%         Conservative:  238
Best Local Similarity:  55.40%         Mismatches:    269
Query Match:            59.12%         Indels:        42
DB:                     6              Gaps:          8

US-09-873-409-6 (1-1195) x AX105078 (1-4279)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
Db    227 GTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy     35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
Db    287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy     45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db    347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy     60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db    407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy     80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db    467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy    100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db    527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy    119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db    587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy    139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db    647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy    159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db    707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy    179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db    767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy    199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
Db    827 GGACAAAAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAGGAATTGGG 886

Qy    219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238
Db    887 ATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCA 946

Qy    239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
Db    947 TCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TAT 1000

Qy    259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
Db   1001 TCTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGA 1060

Qy    279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
Db   1061 CAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTC 1120

Qy    299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
Db   1121 AAGATAATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGAT 1180

Qy    319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
Db   1181 AATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAA 1240

Qy    339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
Db   1241 GTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTT 1300

Qy    359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
Db   1301 GGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCC 1360

Qy    379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
Db   1361 ACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTT 1420

Qy    399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
Db   1421 CGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAA 1480

Qy    419 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
Db   1481 AACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAA 1540

Qy    439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
Db   1541 GCCAATGCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAG 1600

Qy    459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
Db   1601 AGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTT 1660

Qy    479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
Db   1661 CGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAA 1720

Qy    499 SerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 518
Db   1721 GCAGTGGTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCT 1780

Qy    519 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 538
Db   1781 CATCGTTTGTCTACAGTTCGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATT 1840

Qy    539 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 558
Db   1841 GTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTC 1900

Qy    559 MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 578
Db   1901 ACAATGCAG---ACAAGAGGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAA 1957

Qy    579 ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--- 597
Db   1958 AGTGAAAGTGATGCCTTGGAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAA 2014

Qy    598 ----------------------------------------AspLysAlaGluGluSer 603
Db   2015 AGAAGATCAACTCGCAGGAGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACA 2074

Qy    604 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 623
```

```
Db   2075 AAAGAGGACTTGAATGAGAATGTACCTCCAGTTTCCTTCTGGAGGATTCTGAAGCTGAAC 2134

Qy    624 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 643
Db   2135 TCAACTGAATGGCCTTATTTTGTGGTTGGTATATTTTGTGCTATTATAAACGGAGGCCTG 2194

Qy    644 HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAsp 662
Db   2195 CAACCAGCATTTTCAATAATATTTTCAAGGATTATAGGGATCTTTACCCGAGATGAGGAT 2254

Qy    663 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 682
Db   2255 CCTGAAACAAAACGACAGAATAGTAACATGTTTTCTGTATTGTTTCTAGTCCTTGGAATT 2314

Qy    683 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 702
Db   2315 ATTTCTTTTATTACATTTTTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTC 2374

Qy    703 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 722
Db   2375 ACTAAGCGGCTTCGATACATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTT 2434

Qy    723 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 742
Db   2435 GATGACCCTAAAAACACCACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAA 2494

Qy    743 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 762
Db   2495 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGG 2554

Qy    763 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 782
Db   2555 ACAGGCATTATTATATCCTTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATT 2614

Qy    783 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 802
Db   2615 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 2674

Qy    803 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 822
Db   2675 AAAGATAAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTC 2734

Qy    823 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 842
Db   2735 CGAACTGTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTG 2794

Qy    843 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 862
Db   2795 CAAGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATC 2854

Qy    863 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 882
Db   2855 ACCCAGGCAATGATGTATTTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTG 2914

Qy    883 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 902
Db   2915 GCAAATGAGTTCATGAACTTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGT 2974

Qy    903 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 922
Db   2975 GCCATGGCAGTGGGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCA 3034

Qy    923 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 942
Db   3035 GCAGCCCACGTCATCATGATCATTGAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCAC 3094

Qy    943 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 962
Db   3095 GGCCTCAAGCCAAATACGTTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTAT 3154

Qy    963 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 982
Db   3155 CCCACTCGACCAGACATCCCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAG 3214

Qy    983 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 1002
Db   3215 ACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAG 3274

Qy   1003 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 1022
Db   3275 CGCTTCTATGACCCCTTGGCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTG 3334

Qy   1023 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 1042
Db   3335 AATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGAC 3394

Qy   1043 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 1062
Db   3395 TGCAGCATTGCCGAGAACATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAG 3454

Qy   1063 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 1082
Db   3455 ATTATGCAGGCAGCCAAGGAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAA 3514

Qy   1083 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1102
Db   3515 TACAACACCAGAGTAGGAGACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATT 3574

Qy   1103 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 1122
Db   3575 GCCATAGCTCGCGCTCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCTACATCA 3634

Qy   1123 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1142
Db   3635 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 3694

Qy   1143 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1162
Db   3695 CGCACCTGCATTGTGATCGCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTG 3754

Qy   1163 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1182
Db   3755 GTGTTTCAGAATGGCAAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGGCCCAGAAA 3814

Qy   1183 AspIleTyrPheLysLeuValAsnAlaGlnSer 1193
Db   3815 GGCATCTATTTTTCCATGGTCAGTGTCCAGGCT 3847
```

```
RESULT 13
AX105080
LOCUS       AX105080                4279 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 24 from Patent WO0123540.
ACCESSION   AX105080
VERSION     AX105080.1  GI:13921230
KEYWORDS    .
SOURCE      dog.
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 4279)
  AUTHORS   Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and
            Patten,C.J.
  TITLE     P-glycoproteins and uses thereof
  JOURNAL   Patent: WO 0123540-A 24 05-APR-2001;
            GENTEST CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..4279
                     /organism="Canis familiaris"
                     /db_xref="taxon:9615"
     CDS             17..3862
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC37722.1"
                     /db_xref="GI:13921231"
                     /translation="MDPEGGRKGSAEKNFWKMGKKSKKKEKKEKKPTVSTFAMFRYSN
```

```
                    WLDRLYMLVGTMAAIIHGAALPLMMLVFGNMTDSFANAGISRNKTFPVIINESITNNT
                    QHFINHLEEEMTTYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQILKIRKQFFHAIMRQ
                    EIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSIATFFTGFIVGFTRGWKLTL
                    VILAISPVLGLSAAIWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFGGQKKEL
                    ERYNKNLEEAKGIGIKKAITANISIGAAFLLIYASYALAFWYGTSLVLSSEYSIGQVL
                    TVFFSVLIGAFSIGQASPSIEAFANARGAAYEIFKIIDNKPSIDSYSKSGHKPDNIKG
                    NLEFKNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQRLYDPTD
                    GMVCIDGQDIRTINVRHLREITGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKE
                    ANAYDFIMKLPNKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTE
                    SEAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELMKEKGIY
                    FKLVTMQTRGNEIELENATGESKSESDALEMSPKDSGSSLIKRRSTRRSIHAPQGQDR
                    KLGTKEDLNENVPPVSFWRILKLNSTEWPYFVVGIFCAIINGGLQPAFSIIFSRIIGI
                    FTRDEDPETKRQNSNMFSVLFLVLGIISFITFFLQGFTFGKAGEILTKRLRYMVFRSM
                    LRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAVITQNIANLGTGIIISLIYG
                    WQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRTVVSLTR
                    EQKFEYMYAQSLQVPYRNSLRKAHIFGVSFSITQAMMYFSYAGCFRFGAYLVANEFMN
                    FQDVLLVFSAIVFGAMAVGQVSSFAPDYAKAKVSAAHVIMIIEKSPLIDSYSPHGLKP
                    NTLEGNVTFNEVVFNYPTRPDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERF
                    YDPLAGSVLIDGKEIKHLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSHEE
                    IMQAAKEANIHHFIETLPEKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEA
                    TSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGKVKEHGTHQQL
                    LAQKGIYFSMVSVQAGAKR"
BASE COUNT     1296 a    833 c    1008 g    1142 t
ORIGIN

Alignment Scores:
Pred. No.:                5.21e-222      Length:         4279
Score:                    3554.00        Matches:        682
Percent Similarity:       74.74%         Conservative:   238
Best Local Similarity:    55.40%         Mismatches:     269
Query Match:              59.12%         Indels:         42
DB:                       6              Gaps:           8

US-09-873-409-6 (1-1195) x AX105080 (1-4279)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||:::::|||    :::|||:::::::::::||||||    ||||||||||||    |||
Db    167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
           |||   |||   |||:::|||:::       :::       :::
Db    227 GTTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy     35 --------------------ValGlnThrAsnThrTyrSerPhePheArg---------- 44
                              :::      ||||||      |||
Db    287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy     45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                               ||||||   |||||||||          |||:::
Db    347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy     60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
              |||||||||:::|||   |||    :::   |||   ||||||      :::||||||
Db    407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy     80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
           |||||||||||||||:::::::::      |||:::|||||||||||||||     |||:::
Db    467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy    100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
           ||||||||||||||||||:::|||      |||:::   |||::::::|||||||||||||
Db    527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy    119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
           ||||||     :::   ||||||::::::::::||||||      |||      ||||||
Db    587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy    139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
              :::||||||||||||||||||||||||      |||:::   ||||||::::::      |||
Db    647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy    159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
           ||||||      ::::::::::::      |||   |||      |||||||||    |||||||:::
Db    707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy    179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
           ||||||||||||||||||||||||||||||::::::||||||||||||||||||||||||
Db    767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy    199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
              |||:::||||||||:::||||||      :::||||||:::::::||||||      |||
Db    827 GGACAAAAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAGGAATTGGG 886

Qy    219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238
           ||||||:::      |||    :::    :::||||:::||||||         :::
Db    887 ATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCA 946

Qy    239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
           :::|||   ||||||||||||||||||||||||||||:::|||:::      |||      |||
Db    947 TCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TAT 1000

Qy    259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
           :::||||||   ||||||   ||||||||||||||||:::          ::::::   ||||||
Db   1001 TCTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGA 1060

Qy    279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
              |||   |||         |||   ||||||   ||||||||||||||||:::   ||||||
Db   1061 CAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTC 1120

Qy    299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
           ::::::||||||      ||||||||||||||||:::::::|||   |||:::||||||:::
Db   1121 AAGATAATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGAT 1180

Qy    319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
           :::|||:::|||     :::||||||||||||||||    |||:::||||||||||||
Db   1181 AATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAA 1240

Qy    339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
           :::||||||||||||||||||||||||:::::::::||||||:::|||||||||||||||
Db   1241 GTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTT 1300

Qy    359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
           |||    :::|||   ||||||||||||   ||||||||||:::||||||||||||||||||
Db   1301 GGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCC 1360

Qy    379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
           ||||||    :::      :::|||         |||   :::|||||||||||
Db   1361 ACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTT 1420

Qy    399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
           |||:::         ||||||||||||||||||||||||||||   ||||||||:::
Db   1421 CGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAA 1480

Qy    419 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
           ||||||:::|||||||||::::::||||||      :::|||:::|||:::|||      :::|||
Db   1481 AACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAA 1540

Qy    439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
           ||||||||||||||||||||||||:::      ||||||||||||:::||||||||||||||||
Db   1541 GCCAATGCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAG 1600

Qy    459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
           :::|||||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db   1601 AGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTT 1660

Qy    479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
           ||||||||||||||||||:::|||||||||||||||||||||||||||:::||||||:::
Db   1661 CGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAA 1720

Qy    499 SerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 518
           :::   ||||||   ||||||:::||||||   |||||||||||||||||||:::|||
Db   1721 GCAGTGGTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCT 1780

Qy    519 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 538
           ||||||||||||:::|||:::||||||:::|||               ||||||::::::
Db   1781 CATCGTTTGTCTACAGTTCGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATT 1840
```

```
Qy    539 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 558
Db   1841 GTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTC 1900

Qy    559 MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 578
Db   1901 ACAATGCAG---ACAAGAGGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAA 1957

Qy    579 ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--- 597
Db   1958 AGTGAAAGTGATGCCTTGGAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAA 2014

Qy    598 ------------------------------------------AspLysAlaGluGluSer 603
Db   2015 AGAAGATCAACTCGCAGGAGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACA 2074

Qy    604 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 623
Db   2075 AAAGAGGACTTGAATGAGAATGTACCTCCAGTTTCCTTCTGGAGGATTCTGAAGCTGAAC 2134

Qy    624 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 643
Db   2135 TCAACTGAATGGCCTTATTTTGTGGTTGGTATATTTTGTGCTATTATAAACGGAGGCCTG 2194

Qy    644 HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAsp 662
Db   2195 CAACCAGCATTTTCAATAATATTTTCAAGGATTATAGGGATCTTTACCCGAGATGAGGAT 2254

Qy    663 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 682
Db   2255 CCTGAAACAAAACGACAGAATAGTAACATGTTTTCTGTATTGTTCTAGTCCTTGGAATT 2314

Qy    683 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 702
Db   2315 ATTTCTTTTATTACATTTTTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTC 2374

Qy    703 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 722
Db   2375 ACTAAGCGGCTTCGATACATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTT 2434

Qy    723 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 742
Db   2435 GATGACCCTAAAAACACCACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAA 2494

Qy    743 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 762
Db   2495 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGG 2554

Qy    763 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 782
Db   2555 ACAGGCATTATTATATCCTTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATT 2614

Qy    783 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 802
Db   2615 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 2674

Qy    803 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 822
Db   2675 AAAGATAAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTC 2734

Qy    823 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 842
Db   2735 CGAACTGTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTG 2794

Qy    843 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 862
Db   2795 CAAGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATC 2854

Qy    863 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 882
Db   2855 ACCCAGGCAATGATGTATTTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTG 2914

Qy    883 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 902
Db   2915 GCAAATGAGTTCATGAACTTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGT 2974

Qy    903 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 922
Db   2975 GCCATGGCAGTGGGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCA 3034

Qy    923 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 942
Db   3035 GCAGCCCACGTCATCATGATCATTGAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCAC 3094

Qy    943 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 962
Db   3095 GGCCTCAAGCCAAATACGTTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTAT 3154

Qy    963 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 982
Db   3155 CCCACTCGACCAGACATCCCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAG 3214

Qy    983 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 1002
Db   3215 ACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAG 3274

Qy   1003 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 1022
Db   3275 CGCTTCTATGACCCCTTGGCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTG 3334

Qy   1023 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 1042
Db   3335 AATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGAC 3394

Qy   1043 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 1062
Db   3395 TGCAGCATTGCCGAGAACATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAG 3454

Qy   1063 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 1082
Db   3455 ATTATGCAGGCAGCCAAGGAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAA 3514

Qy   1083 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1102
Db   3515 TACAACACCAGAGTAGGAGACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATT 3574

Qy   1103 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 1122
Db   3575 GCCATAGCTCGCGCTCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCTACATCA 3634

Qy   1123 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1142
Db   3635 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 3694

Qy   1143 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1162
Db   3695 CGCACCTGCATTGTGATCGCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTG 3754

Qy   1163 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1182
Db   3755 GTGTTTCAGAATGGCAAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGGCCCAGAAA 3814

Qy   1183 AspIleTyrPheLysLeuValAsnAlaGlnSer 1193
Db   3815 GGCATCTATTTTTCCATGGTCAGTGTCCAGGCT 3847
```

```
RESULT 14
AX391099
LOCUS       AX391099              4646 bp    DNA     linear   PAT 19-MAR-2002
DEFINITION  Sequence 1 from Patent WO0210205.
ACCESSION   AX391099
VERSION     AX391099.1  GI:19584228
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Ruth,A. and Roninson,I.
  TITLE     Mutations of the mdr1 p-glycoprotein that improve its ability to
            confer resistance to chemotherapeutic drugs
  JOURNAL   Patent: WO 0210205-A 1 07-FEB-2002;
            THE BOARD OF TRUSTEES OF THE UNIVERSITY OF ILLINOIS (US)
FEATURES             Location/Qualifiers
     source          1..4646
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     CDS             425..4267
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAD28599.1"
                     /db_xref="GI:19584229"
                     /translation="MDLEGDRNGGAKKKNFFKLNNKSEKDKKEKKPTVSVFSMFRYSN
                     WLDKLYMVVGTLAAIIHGAGLPLMMLVFGEMTDIFANAGNLEDLMSNITNRSDINDTG
                     FFMNLEEDMTRYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQIHKIRKQFFHAIMRQEI
                     GWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSMATFFTGFIVGFTRGWKLTLVI
                     LAISPVLGLSAAVWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFGGQKKELER
                     YNKNLEEAKRIGIKKAITANISIGAAFLLIYASYALAFWYGTTLVLSGEYSIGQVLTV
                     FFSVLIGAFSVGQASPSIEAFANARGAAYEIFKIIDNKPSIDSYSKSGHKPDNIKGNL
                     EFRNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQRLYDPTEGM
                     VSVDGQDIRTINVRFLREIIGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKEAN
                     AYDFIMKLPHKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTESE
                     AVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELMKEKGIYFK
                     LVTMQTAGNEVELENAADESKSEIDALEMSSNDSRSSLIRKRSTRRSVRGSQAQDRKL
                     STKEALDESIPPVSFWRIMKLNLTEWPYFVVGVFCAIINGGLQPAFAIIFSKIIGVFT
                     RIDDPETKRQNSNLFSLLFLALGIISFITFFLQGFTFGKAGEILTKRLRYMVFRSMLR
                     QDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAVITQNIANLGTGIIISFIYGWQ
                     LTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRTVVSLTQEQ
                     KFEHMYAQSLQVPYRNSLRKAHIFGITFSFTQAMMYFSYAGCFRFGAYLVAHKLMSFE
                     DVLLVFSAVVFGAMAVGQVSSFAPDYAKAKISAAHIIMIIEKTPLIDSYSTEGLMPNT
                     LEGNVTFGEVVFNYPTRPDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERFYD
                     PLAGKVLLDGKEIKRLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSQEEIV
                     RAAKEANIHAFIESLPNKYSTKVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEATS
                     ALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGRVKEHGTHQQLLA
                     QKGIYFSMVSVQAGTKRQ"
BASE COUNT     1372 a    892 c    1130 g    1252 t
ORIGIN

Alignment Scores:
Pred. No.:              5.78e-222       Length:         4646
Score:                  3554.00         Matches:        682
Percent Similarity:     75.63%          Conservative:   246
Best Local Similarity:  55.58%          Mismatches:     263
Query Match:            59.12%          Indels:         36
DB:                     6               Gaps:           9

US-09-873-409-6 (1-1195) x AX391099 (1-4646)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            ||:::::|||   ||||||:::::::::::|||||   |||||||||||    |||
Db      575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy       21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
            |||   |||||||||:::|||                           :::|||:::|||
Db      635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy       32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                     ::::::      |||          |||||
Db      695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy       44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                  |||         ||||||   ||||||||      |||:::          |||
Db      755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy       62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
            ||||||:::|||   |||   :::   |||   |||      :::||||||||||||
Db      815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy       82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
            |||||||||:::::::::      |||:::|||||||||||||||      |||:::||||||
Db      875 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy      102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
            ||||||||||||:::|||   |||:::   |||||||:::::::||||||||||||||||
Db      935 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGGAATTGGTGACAAAATT 994

Qy      121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
               :::   ||||||:::|||:::||||||         |||         ||||||         :::
Db      995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy      141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
            ||||||||||||||||||||||      |||:::      ||||||::::::      |||||||||
Db     1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy      161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
                ::::::::::::      |||      |||      |||||||||      ||||||:::||||||
Db     1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy      181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
            |||||||||||||||||||||||:::::::|||||||||||||||||||||         |||
Db     1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy      201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
            :::|||||||||:::||||||      :::||||||:::::::||||||         |||||||||
Db     1235 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1294

Qy      221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
            :::   |||   :::      :::|||:::||||||      :::         :::         :::|||
Db     1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1354

Qy      241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
               ||||||||||||||||||||||:::|||:::|||:::||||||         |||:::|||
Db     1355 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1408

Qy      261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
            |||   ||||||      ||||||||||||||||:::         ::::::      :::|||   |||
Db     1409 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1468

Qy      281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
               |||         |||      ||||||   |||||||||||||||:::      ||||||::::::
Db     1469 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528

Qy      301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
            ||||||   |||||||||||||||:::::::|||      :::|||:::||||||:::::::|||
Db     1529 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1588

Qy      321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
            :::|||      :::|||||||:::||||||      |||:::||||||||||||         :::|||
Db     1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1648

Qy      341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
            |||||||||||||||||||||||:::::::||||||:::||||||||||||||||||
Db     1649 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1708

Qy      361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
            :::|||   ||||||||||||   ||||||||||:::|||||||||||||||||||         :::
Db     1709 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1768

Qy      381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
            |||   :::   ||||||      |||||||||      :::|||||||||      |||:::
Db     1769 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1828

Qy      401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
               ||||||||||||||||||||||||||||||   ||||||||||:::      ||||||
Db     1829 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1888

Qy      421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
            :::||||||||::::::||||||      :::|||:::|||:::|||      :::||||||||
Db     1889 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1948

Qy      441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
```

```
                |||||||||||||||||:::    |||:::||||||:::||||||||||||||:::|||
Db       1949 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2008

Qy        461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
              ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       2009 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2068

Qy        481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
              ||||||||||||:::||||||||||||||||||||||||||||:::||||||||::::::
Db       2069 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2128

Qy        501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
              ||||||   ||||||:::||||||   ||||||||||||||||||||:::||||||||||
Db       2129 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 2188

Qy        521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
              |||||||||:::|||:::||||||:::|||            ||||||::::::   |||
Db       2189 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 2248

Qy        541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
              ||||||   |||   |||||||||      ::::::|||:::|||:::   ||||||
Db       2249 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 2308

Qy        561 Gln------------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
              |||               :::::::: |||   ::::::   :::|||
Db       2309 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 2368

Qy        575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 589
                 :::   |||   :::|||               :::
Db       2369 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2428

Qy        590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
                 :::|||::::::   |||         |||:::         :::   ::::::
Db       2429 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2488

Qy        608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627
              :::   |||:::|||   ||||||      :::|||:::|||||||||      ||||||
Db       2489 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2548

Qy        628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
              |||:::   |||:::|||         :::::::::|||||   :::   |||   |||
Db       2549 CCTTATTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2608

Qy        648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
              :::||||||||||:::|||||||||      :::|||      :::|||      |||
Db       2609 GCAATAATATTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2668

Qy        667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
              :::   ::::::   ::::::|||::::::|||:::   ||||||:::|||   |||:::
Db       2669 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2728

Qy        687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
              ::::::|||:::||||||         :::|||:::||||||||||||||||   ||||||
Db       2729 ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2788

Qy        707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
              |||::::::   |||::::::||||||   ||||||::::::|||||||||:::   :::
Db       2789 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2848

Qy        727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
              |||:::||||||   |||||||||   ||||||   |||   ||||||::::::||||||
Db       2849 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2908

Qy        747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
                 |||||||||:::   |||:::|||||||||      |||:::|||      :::|||
Db       2909 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2968

Qy        767 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
              ||||||||||||||||||||::::::|||   |||:::|||:::|||   |||::::::
Db       2969 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 3028

Qy        787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
              |||:::   |||::::::|||            ::::::|||   |||   ||||||:::
Db       3029 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 3088

Qy        807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
              ||||||:::   |||||||||||||||||||||||:::||||||   ||||||:::|||
Db       3089 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 3148

Qy        827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
              |||||||||:::|||:::      ||||||   ||||||   :::   ||||||   :::
Db       3149 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 3208

Qy        847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
              ||||||:::   :::||||||   |||   |||      ::::::|||:::   |||
Db       3209 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 3268

Qy        867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
              :::||||||:::||||||      |||||||||||||||||||||:::
Db       3269 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 3328

Qy        887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
              |||:::   |||   :::   :::||||||:::|||:::   :::||||||||||||:::
Db       3329 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 3388

Qy        907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
              |||:::               ||||||:::|||:::|||||||||   |||||||||:::
Db       3389 GGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3448

Qy        927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
                 ::::::||||||   |||   |||||||||   |||   ||||||      |||
Db       3449 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3508

Qy        947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
              :::|||   |||||||||:::   |||   ||||||   |||   ||||||   ||||||
Db       3509 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3568

Qy        967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
              |||:::   :::|||:::||||||||||||   ::::::::::|||:::|||:::|||
Db       3569 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3628

Qy        987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
              ||||||||||||||||||||||||||||||   ||||||||||||:::|||   ||||||
Db       3629 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3688

Qy       1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
              |||:::   |||:::||||||   ||||||   :::   |||   |||||||||||||||
Db       3689 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3748

Qy       1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
              ||||||:::   :::   ||||||   |||||||||:::||||||:::||||||||||||
Db       3749 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3808

Qy       1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
              ||||||||||||||||||||||||||||||||||||         :::||||||   |||
Db       3809 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3868

Qy       1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
              |||      ||||||||||||:::|||||||||   ||||||   ||||||:::|||:::
Db       3869 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3928

Qy       1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
              ||||||   ||||||   |||||||||||||||||||||||||:::||||||||||||
Db       3929 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3988

Qy       1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
              ||||||:::::::::|||   |||||||||||||||||||||||||||||||||
Db       3989 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 4048

Qy       1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
              :::||||||||||||||||||   ||||||||||||||||||   ||||||||||||:::
Db       4049 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 4108

Qy       1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
              |||:::   ||||||||||||   |||||||||||||||||||||||||||      |||
Db       4109 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 4168
```

```
Qy    1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
           |||::::::||||||   ||||||||||||:::||||||       :::    ||||||||
Db    4169 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 4228

Qy    1187 LysLeuValAsnAlaGlnSer 1193
              :::|||:::   |||:::
Db    4229 TCAATGGTCAGTGTCCAGGCT 4249

RESULT 15
AF269224
LOCUS       AF269224                4045 bp    DNA     linear   SYN 04-JUL-2000
DEFINITION  Synthetic construct Canis familiaris his-tagged-multidrug
            resistance glycoprotein gene, complete cds.
ACCESSION   AF269224
VERSION     AF269224.1  GI:8926216
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4045)
  AUTHORS   Roulet,A.C., Puel,O., Gesta,S., Alvinerie,M. and Pineau,T.J.
  TITLE     Analysis of dog MDR1 p-glycoprotein
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4045)
  AUTHORS   Roulet,A.C., Puel,O., Gesta,S., Alvinerie,M. and Pineau,T.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-MAY-2000) Pharmacology, INRA, 180 Chemin de
            Tournefeuille, BP3, Toulouse 31931, France
FEATURES             Location/Qualifiers
     source          1..4045
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="original mRNA isolated from Beagle dog colon"
     CDS             101..3964
                     /codon_start=1
                     /transl_table=11
                     /product="his-tagged-multidrug resistance glycoprotein
                     MDR1"
                     /protein_id="AAF81747.1"
                     /db_xref="GI:8926217"
                     /translation="MHHHHHHDPEGGRKGSAEKNFWKMGKKSKKNEKKEKKPTVSTFA
                     MFRYSNWLDRLYMLVGTMAAIIHGAALPLMMLVFGNMTDSFANAGISRNKTFPVIINE
                     SITNNTQHFINHLEEEMTTYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQILKIRKQFF
                     HAIMRQEIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSIATFFTGFIVGFTR
                     GWKLTLVILAISPVLGLSAAIWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFG
                     GQKKELERYNKNLEEAKRIGIKKAITANISIGAAFLLIYASYALAFWYGTSLVLSSEY
                     SIGQVLTVFFSVLIGAFSIGQASPSIEAFANARGAAYEIFKIIDNKPSIDSYSKSGHK
                     PDNIKGNLEFKNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQR
                     LYDPTDGMVCIDGQDIRTINVRHLREITGVVSQEPVLFATTIAENIRYGRENVTMDEI
                     EKAVKEANAYDFIMKLPNKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEAT
                     SALDTESEAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELM
                     KEKGIYFKLVTMQTRGNEIELENATGESKSESDALEMSPKDSGSSLIKRRSTRRSIHA
                     PQGQDRKLGTKEDLNENVPPVSFWRILKLNSTEWPYFVVGIFCAIINGGLQPAFSIIF
                     SRIIGIFTRDEDPETKRQNSNMFSVLFLVLGIISFITFFLQGFTFGKAGEILTKRLRY
                     MVFRSMLRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAVITQNIANLGTGII
                     ISLIYGWQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRT
                     VVSLTREQKFEYMYAQSLQVPYRNSLRKAHIFGVSFSITQAMMYFSYAGCFRFGAYLV
                     ANEFMNFQDVLLVFSAIVFGAMAVGQVSSFAPDYAKAKVSAAHVIMIIEKSPLIDSYS
                     PHGLKPNTLEGNVTFNEVVFNYPTRPDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVV
                     QLLERFYDPLAGSVLIDGKEIKHLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSR
                     VVSHEEIMQAAKEANIHHFIETLPEKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHI
                     LLLDEATSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGKVKEH
                     GTHQQLLAQKGIYFSMISVQAGAKR"
BASE COUNT     1171 a    850 c    986 g   1038 t
ORIGIN

Alignment Scores:
Pred. No.:              6.54e-222       Length:         4045
Score:                  3552.00         Matches:        681
Percent Similarity:     74.74%          Conservative:   239
Best Local Similarity:  55.32%          Mismatches:     269
Query Match:            59.08%          Indels:         42
DB:                          12         Gaps:           8

US-09-873-409-6 (1-1195) x AF269224 (1-4045)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||::::::|||    :::|||:::::::::::::|||||    ||||||||||    |||
Db     269 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 328

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
           |||   |||   |||:::|||:::       :::       :::
Db     329 GTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 388

Qy      35 --------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                               :::       ||||||      |||
Db     389 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 448

Qy      45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                          ||||||   |||||||||          |||:::
Db     449 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 508

Qy      60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
              ||||||||:::|||   |||   :::   |||   ||||||       :::||||||
Db     509 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 568

Qy      80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
           ||||||||||||||||:::::::::   |||:::|||||||||||||||      |||:::
Db     569 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 628

Qy     100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
           ||||||||||||||||||:::|||   |||:::   ||||||::::::||||||||||||
Db     629 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 688

Qy     119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
           ||||||   :::   ||||||:::::::::||||||      |||      ||||||
Db     689 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 748

Qy     139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
           :::||||||||||||||||||||||||      |||   ||||||::::::      |||
Db     749 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 808

Qy     159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
           ||||||   ::::::::::::   |||   |||   |||||||||   ||||||:::
Db     809 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 868

Qy     179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
           ||||||||||||||||||||||||||||||::::::||||||||||||||||||||||
Db     869 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 928

Qy     199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
              |||:::||||||||:::||||||   :::||||||::::::||||||      |||
Db     929 GGACAAAAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGG 988

Qy     219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238
           ||||||:::   |||   :::   :::|||:::||||||       :::       :::
Db     989 ATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCA 1048

Qy     239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
           :::|||   ||||||||||||||||||||||||||||:::|||:::   |||      |||
Db    1049 TCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TAT 1102

Qy     259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
           :::||||||   ||||||   |||||||||||||||:::      ::::::   ||||||
Db    1103 TCTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGA 1162

Qy     279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
              |||   |||      |||   ||||||   ||||||||||||||||:::   ||||||
Db    1163 CAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTC 1222

Qy     299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
           ::::::||||||   ||||||||||||||||||:::::|||   :::|||:::||||||:::
Db    1223 AAGATAATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGAT 1282

Qy     319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
```

```
Db   1283 AATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAA 1342
Qy    339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
Db   1343 GTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTT 1402
Qy    359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
Db   1403 GGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCC 1462
Qy    379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
Db   1463 ACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTT 1522
Qy    399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
Db   1523 CGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAA 1582
Qy    419 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
Db   1583 AACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAA 1642
Qy    439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
Db   1643 GCCAATGCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAG 1702
Qy    459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
Db   1703 AGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTT 1762
Qy    479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
Db   1763 CGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAA 1822
Qy    499 SerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 518
Db   1823 GCAGTGGTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCT 1882
Qy    519 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 538
Db   1883 CATCGTTTGTCTACAGTTCGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATT 1942
Qy    539 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 558
Db   1943 GTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTC 2002
Qy    559 MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 578
Db   2003 ACAATGCAG---ACAAGAGGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAA 2059
Qy    579 ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--- 597
Db   2060 AGTGAAAGTGATGCCTTGGAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAA 2116
Qy    598 ------------------------------------------AspLysAlaGluGluSer 603
Db   2117 AGAAGATCAACTCGCAGGAGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACA 2176
Qy    604 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 623
Db   2177 AAAGAGGACTTGAATGAGAATGTACCTCCAGTTTCCTTCTGGAGGATTCTGAAGCTGAAC 2236
Qy    624 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 643
Db   2237 TCAACTGAATGGCCTTATTTTGTGGTTGGTATATTTTGTGCTATTATAAACGGAGGCCTG 2296
Qy    644 HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAsp 662
Db   2297 CAACCAGCATTTTCAATAATATTTTCAAGGATTATAGGGATCTTTACCCGAGATGAGGAT 2356
Qy    663 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 682
Db   2357 CCTGAAACAAAACGACAGAATAGTAACATGTTTTCTGTATTGTTTCTAGTCCTTGGAATT 2416
Qy    683 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 702
Db   2417 ATTTCTTTTATTACATTTTTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTC 2476
Qy    703 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 722
Db   2477 ACTAAGCGGCTTCGATACATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTT 2536
Qy    723 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 742
Db   2537 GATGACCCTAAAAACACCACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAA 2596
Qy    743 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 762
Db   2597 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGG 2656
Qy    763 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 782
Db   2657 ACAGGCATTATTATATCCTTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATT 2716
Qy    783 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 802
Db   2717 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 2776
Qy    803 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 822
Db   2777 AAAGATAAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTC 2836
Qy    823 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 842
Db   2837 CGAACTGTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTG 2896
Qy    843 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 862
Db   2897 CAAGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATC 2956
Qy    863 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 882
Db   2957 ACCCAGGCAATGATGTATTTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTG 3016
Qy    883 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 902
Db   3017 GCAAATGAGTTCATGAACTTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGT 3076
Qy    903 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 922
Db   3077 GCCATGGCAGTGGGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCA 3136
Qy    923 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 942
Db   3137 GCAGCCCACGTCATCATGATCATTGAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCAC 3196
Qy    943 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 962
Db   3197 GGCCTCAAGCCAAATACGTTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTAT 3256
Qy    963 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 982
Db   3257 CCCACTCGACCAGACATCCCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAG 3316
Qy    983 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 1002
Db   3317 ACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAG 3376
Qy   1003 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 1022
Db   3377 CGCTTCTATGACCCCTTGGCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTG 3436
Qy   1023 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 1042
Db   3437 AATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGAC 3496
```

```
Qy    1043 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 1062
           |||||||||||||||||||||||||||||||||||||||||||||||||||||       :::|||
Db    3497 TGCAGCATTGCCGAGAACATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAG 3556

Qy    1063 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 1082
           |||   :::||||||      ||||||||||||   |||||||||   ||||||||||||
Db    3557 ATTATGCAGGCAGCCAAGGAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAA 3616

Qy    1083 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1102
           |||||||||:::||||||   ||||||   |||||||||||||||||||||||||||:::
Db    3617 TACAACACCAGAGTAGGAGACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATT 3676

Qy    1103 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 1122
           ||||||||||||||||||:::::::::|||   |||||||||||||||||||||||||||
Db    3677 GCCATAGCTCGCGCTCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCTACATCA 3736

Qy    1123 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1142
           |||||||||   :::||||||||||||||||||   ||||||||||||||||||   |||
Db    3737 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 3796

Qy    1143 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1162
           |||||||||:::|||:::   ||||||||||||   ||||||||||||||||||||||||
Db    3797 CGCACCTGCATTGTGATCGCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTG 3856

Qy    1163 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1182
           |||      ||||||||||:::||||||   ||||||||||||:::||||||      :::
Db    3857 GTGTTTCAGAATGGCAAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGGCTCAGAAA 3916

Qy    1183 AspIleTyrPheLysLeuValAsnAlaGlnSer 1193
              |||||||||   :::::::::   |||:::
Db    3917 GGCATCTATTTTTCCATGATCAGTGTCCAGGCT 3949
```

Search completed: March 31, 2003, 02:22:27
Job time : 16406.1 secs

```
                    GenCore version 5.1.4_p5_4578
          Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 30, 2003, 01:20:01 ; Search time 1167.91 Seconds
                                          (without alignments)
                                          2304.231 Million cell updates/sec

Title:          US-09-873-409-6
Perfect score:  6012
Sequence:       1 MILGILASLVNGACLPLMPL..........QELLRNRDIYFKLVNAQSVQ 1195

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters:       4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115416_19216/app_query.fasta_1.7544
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_4247_@runat_27032003_115416_19216 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      N_Geneseq_101002:*
                1:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
                2:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
                3:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
                4:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
                5:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
                6:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
                7:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
                8:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
                9:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
                10:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
                11:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
                12:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
                13:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
                14:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
                15:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
                16:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
                17:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
                18:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
                19:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
                20:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
                21:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
                22:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
                23:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
                24:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5760.5	95.8	3699	24	ABK83223	Human transporter
2	3557	59.2	4279	22	AAD03506	Dog P-glycoprotein
3	3556.5	59.2	4195	22	AAF86128	Cynomologous monke
4	3554	59.1	3860	21	AAZ49332	Human wild-type mu
5	3554	59.1	3860	24	ABA94365	Human BCRP DNA rel
6	3554	59.1	4279	22	AAD03504	Dog P-glycoprotein
7	3554	59.1	4279	22	AAD03505	Dog P-glycoprotein
8	3549	59.0	4279	22	AAD03488	Dog P-glycoprotein
9	3549	59.0	4317	22	AAD03489	Dog P-glycoprotein
10	3548	59.0	4186	22	AAF86127	Cynomologous monke
11	3545	59.0	3860	21	AAZ49333	Human G185V mutant
12	3545	59.0	3860	24	ABA94366	Human BCRP DNA rel
13	3545	59.0	4349	22	AAH57442	Human intestine ce
14	3545	59.0	4646	21	AAZ94738	Human ATP binding
15	3545	59.0	4646	24	AAD38994	Human mdr1 gene.
16	3545	59.0	4646	24	ABL68592	Kidney cancer rela
17	3545	59.0	4646	24	ABL68880	Kidney cancer rela
18	3545	59.0	6505	17	AAT13394	Hybrid vector pSF-
19	3545	59.0	8630	21	AAZ24041	Retroviral M4 mdr-
20	3545	59.0	8630	21	AAZ24042	Retroviral vector
21	3543.5	58.9	4264	19	AAV66533	Mutated human P-gl
22	3543.5	58.9	4264	19	AAV66534	Mutated human P-gl
23	3543	58.9	4669	8	AAN70752	Sequence of human
24	3540.5	58.9	4425	21	AAZ52048	Rat multidrug resi
25	3539	58.9	4378	11	AAQ04522	Multidrug Resistan
26	3538.5	58.9	4369	21	AAZ52047	Rat multidrug resi
27	3537.5	58.8	3840	24	ABL91687	Human polynucleoti
28	3537.5	58.8	3988	21	AAZ88973	Human MDR-1 DNA.
29	3536	58.8	4669	14	AAQ52726	Sequence of human
30	3534	58.8	4646	15	AAQ72872	Human multidrug re
31	3527	58.7	4669	19	AAV32645	Human P glycoprote
32	3527	58.7	4669	24	ABK52041	cDNA encoding huma
33	3526.5	58.7	4189	21	AAZ49334	Murine multidrug r
34	3526.5	58.7	4189	24	ABA94367	Mouse BCRP DNA rel
35	3526.5	58.7	4313	14	AAQ38950	Mouse multidrug re
36	3522.5	58.6	4788	21	AAZ49335	Murine multidrug r
37	3522.5	58.6	4788	24	ABA94368	Mouse BCRP DNA rel
38	3517.5	58.5	4233	21	AAZ90198	Rat mdr1b2 (multis
39	3517.5	58.5	4233	22	AAF27498	Rat mdr1b2 multidr
40	3504.5	58.3	3924	21	AAZ94742	Human ATP binding
41	3504.5	58.3	3924	21	AAZ88974	Human MDR-3 DNA.
42	3504.5	58.3	3924	24	ABN95801	Gene #2299 used to
43	3476.5	57.8	4254	24	ABK63517	Rat sequence diffe
44	3428.5	57.0	3912	24	ABK63653	Rat sequence diffe
45	2868	47.7	4776	21	AAZ94744	Human ATP binding

ALIGNMENTS

```
RESULT 1
ABK83223
ID   ABK83223 standard; cDNA; 3699 BP.
XX
AC   ABK83223;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human transporter and ion channel, TRICH14, Incyte ID 7472030CB1, cDNA.
XX
KW   Human; ss; gene; transporter and ion channel; TRICH; transport disorder;
KW   neurological disorder; muscle disorder; immunological disorder; cancer;
KW   scleroderma; systemic lupus erythematosus; allergy; leukaemia;
KW   cell proliferative disorder; cervical cancer; breast cancer;
KW   neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;
KW   myotonic dystrophy; catatonia; endocrine disorder; diabetes;
KW   Grave's disease; gastrointestinal disorder; Crohn's disease;
KW   renal disorder; Good pasture's syndrome; viral infection; cirrhosis;
```

```
KW   bacterial infection; fungal infection; parasitic infection;
KW   protozoal infection; helminthic infection; cardiovascular disorder;
KW   atherosclerosis; hepatic disease.
XX
OS   Homo sapiens.
XX
PN   WO200240541-A2.
XX
PD   23-MAY-2002.
XX
PF   25-OCT-2001; 2001WO-US46055.
XX
PR   27-OCT-2000; 2000US-243989P.
PR   03-NOV-2000; 2000US-245904P.
PR   09-NOV-2000; 2000US-247673P.
PR   17-NOV-2000; 2000US-249661P.
PR   20-NOV-2000; 2000US-252232P.
PR   01-DEC-2000; 2000US-250790P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Tang YT,  Yue H,  Nguyen DB,  Hafalia AJA,  Elliott VS,  Lu Y;
PI   Walia NK,  Yao MG,  Baughn MR,  Gandhi AR,  Ding L,  Sanjanwala M;
PI   Ramkumar J,  Arvizu C,  Gietzen KJ,  Lal PG,  Azimzai Y,  Khan FA;
PI   Thangavelu K,  Thornton M,  Lu DAM,  Tribouley CM,  Warren BA;
PI   Ison CH,  Das D,  Raumann BE,  Policky JL,  Kearney L;
XX
DR   WPI; 2002-463570/49.
DR   P-PSDB; ABG61544.
XX
PT   New transporters and ion channels (TRICH) polypeptides, useful for
PT   diagnosing, preventing, and treating disorders associated with an
PT   abnormal expression or activity of TRICH, e.g. immunological, muscular
PT   or renal disorders  -
XX
PS   Claim 5; Page 172; 178pp; English.
XX
CC   The invention relates to human transporters and ion channels (TRICH)
CC   polypeptides, a naturally occurring amino acid sequence 90 % identical to
CC   TRICH, a biologically active fragment of TRICH or an immunogenic fragment
CC   of TRICH. Also included are an isolated polynucleotide encoding TRICH,
CC   a recombinant polynucleotide comprising a promoter sequence operably
CC   linked to the TRICH polynucleotide, a cell transformed with the
CC   recombinant polynucleotide, a transgenic organism comprising the
CC   recombinant polynucleotide, an isolated antibody that binds specifically
CC   to TRICH, and screening for compounds which bind to TRICH, modulate
CC   TRICH, modulate TRICH expression or are ant/agonists of TRICH.
CC   The polypeptides are useful for diagnosing, treating, and
CC   preventing transport, neurological, muscle, immunological disorders
CC   (e.g. scleroderma, systemic lupus erythematosus, allergies), cell
CC   proliferative disorders such as cancers (e.g. leukaemia, cervical or
CC   breast cancers), neurodegenerative disorders (e.g. Parkinson's disease,
CC   Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy,
CC   catatonia), endocrine disorders (e.g. diabetes, Grave's disease),
CC   gastrointestinal disorders (e.g. Crohn's disease), renal disorders
CC   (e.g. Good pasture's syndrome), viral, bacterial, fungal, parasitic,
CC   protozoal and helminthic infections, cardiovascular disorders (e.g.
CC   atherosclerosis), or hepatic diseases (e.g. cirrhosis) and many
CC   other diseases and disorders detailed in the specification. They can also
CC   be used in assessing the effects of exogenous compounds on the
CC   expression of nucleic acid and amino acid sequences of transporters and
CC   ion channels. TRICH or its fragments may also be used in screening for
CC   compounds that specifically bind to and modulate the activity of TRICH.
CC   The polynucleotides can be used to create knock-in humanised animals or
CC   transgenic animals to model human disease. The present sequence
CC   encodes a TRICH protein.
XX
SQ   Sequence 3699 BP; 1116 A; 707 C; 860 G; 1016 T; 0 other;

Alignment Scores:
Pred. No.:            0          Length:        3699
Score:                5760.50    Matches:       1156
Percent Similarity:   95.87%     Conservative:  4
```

```
Best Local Similarity:  95.54%        Mismatches:    9
Query Match:            95.82%        Indels:        41
DB:                     24            Gaps:          3

US-09-873-409-6 (1-1195) x ABK83223 (1-3699)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           ||||||||||||||||||   ||||||   ||||||||||||||||||||||||||||||
Db     145 ATGATCCTGGGTATACTGACATCACTGTTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 204

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
              :::|||||||||||||||||||||||||||||||||||||||   |||||||||
Db     205 TGTATAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCACACTAACACAACA 264

Qy      41 SerPhePheArg------------------------------------------LeuThr 46
           ::::::                                                ||||||
Db     265 AATTATCAGAACTGTACTCAGTCTCAAGAGAAGCTGAATGAAGATATGACTCTGTTGACC 324

Qy      47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     325 CTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGTTACATACAGATTTCCTTG 384

Qy      67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     385 TGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAACAGTTTTTTCATTCAGTT 444

Qy      87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     445 TTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGTGAACTTAACACTCGCATG 504

Qy     107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
           |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     505 ACAGATGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATTGCTCTGTTGTTTCAA 564

Qy     126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     565 AACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAGGGCTGGAAACTCACC 624

Qy     146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     625 CTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCAGCATGTTCTAGGATG 684

Qy     166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     685 GTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCAGAA 744

Qy     186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     745 GAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAGGAGAAAGAACTTCAA 804

Qy     206 ArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSer 225
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     805 AGGTATACACAGAATCTCAAAGATGCAAAGGATTTTGGCATAAAAAGGACTATAGCTTCA 864

Qy     226 LysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrp 245
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     865 AAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAATGGAACCTATGGACTTGCTTTTTGG 924

Qy     246 TyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAla 265
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     925 TATGGAACCTCCTTGATTCTTAATGGAGAACCTGGATATACCATCGGGACTGTTCTTGCT 984

Qy     266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     985 GTTTTCTTTAGTGTAATCCATAGTAGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAA 1044

Qy     286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1045 ACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTTATTGATAAGAAACCC 1104

Qy     306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
           ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1105 AGTATAGGTAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGAA 1164
```

```
Qy     326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1165 TTTAAAAATGTTTCTTTCAATTATCCATCAAGGCCATCTATCAAGATTCTGAAAGGTCTG 1224

Qy     346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
           ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1225 AATCTCGGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAG 1284

Qy     366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1285 AGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGTG 1344

Qy     386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1345 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTGGTT 1404

Qy     406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1405 AGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGAGAT 1464

Qy     426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1465 GATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATC 1524

Qy     446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1525 ATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGA 1584

Qy     466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1585 GGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATT 1644

Qy     486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeu 505
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1645 TTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTG 1704

Qy     506 GluLysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArg 525
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1705 GAGAAGGCGAGCAAAGGTCGGACTACAATCGTGGTAGCACACCGACTTTCTACTATTCGA 1764

Qy     526 SerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAla 545
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1765 AGTGCAGATTTGATTGTGACCCTAAAGGATGGAATGCTGGCGGAGAAAGGAGCGCATGCT 1824

Qy     546 GluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGlnAspIleLysLys 565
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1825 GAACTAATGGCAAAACGAGGTCTATATTATTCACTTGTGATGTCACAGGATATTAAAAAA 1884

Qy     566 AlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeuPro 585
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1885 GCTGATGAACAGATGGAGTCAATGACATATTCTACTGAAAGAAAGACCAACTCACTTCCT 1944

Qy     586 LeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThrGln 605
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1945 CTGCACTCTGTGAAGAGCATCAAGTCAGACTTCATTGACAAGGCTGAGGAATCCACCCAA 2004

Qy     606 SerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysPro 625
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2005 TCTAAAGAGATAAGTCTTCCTGAAGTCTCTCTATTAAAAATTTTAAAGTTAAACAAGCCT 2064

Qy     626 GluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisPro 645
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2065 GAATGGCCTTTTGTGGTTCTGGGGACATTGGCTTCTGTTCTAAATGGAACTGTTCATCCA 2124

Qy     646 ValPheSerIleIlePheAlaLysIleIleThrMetPheGlyAsnAsnAspLysThrThr 665
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2125 GTATTTTCCATCATCTTTGCAAAAATTATAACCATGTTTGGAAATAATGATAAAACCACA 2184

Qy     666 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 685
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2185 TTAAAGCATGATGCAGAAATTTATTCCATGATATTCGTCATTTTGGGTGTTATTTGCTTT 2244

Qy     686 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 705
           |||||||||||||||
Db    2245 GTCAGTTATTTCATG--------------------------------------------- 2259

Qy     706 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLys 725
                                                |||||||||||||||||||||||||||
Db    2260 ---------------------------------CAGGATATTGCCTGGTTTGATGAAAAG 2286

Qy     726 GluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 745
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2287 GAAAACAGCACAGGAGGCTTGACAACAATATTAGCCATAGATATAGCACAAATTCAAGGA 2346

Qy     746 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 765
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2347 GCAACAGGTTCCAGGATTGGCGTCTTAACACAAAATGCAACTAACATGGGACTTTCAGTT 2406

Qy     766 IleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVal 785
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2407 ATCATTTCCTTTATATATGGATGGGAGATGACATTCCTGATTCTGAGTATTGCTCCAGTA 2466

Qy     786 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 805
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2467 CTTGCCGTGACAGGAATGATTGAAACCGCAGCAATGACTGGATTTGCCAACAAAGATAAG 2526

Qy     806 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIle 825
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2527 CAAGAACTTAAGCATGCTGGAAAGATAGCAACTGAAGCTTTGGAGAATATACGTACTATA 2586

Qy     826 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGln 845
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2587 GTGTCATTAACAAGGGAAAAAGCCTTCGAGCAAATGTATGAAGAGATGCTTCAGACTCAA 2646

Qy     846 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla 865
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2647 CACAGAAATACCTCGAAGAAAGCACAGATTATTGGAAGCTGTTATGCATTCAGCCATGCC 2706

Qy     866 PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGly 885
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2707 TTTATATATTTTGCCTATGCGGCAGGGTTTCGATTTGGAGCCTATTTAATTCAAGCTGGA 2766

Qy     886 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAla 905
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2767 CGAATGACCCCAGAGGGCATGTTCATAGTTTTTACTGCAATTGCATATGGAGCTATGGCC 2826

Qy     906 IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHis 925
           ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2827 ATCGGAGAAACGCTCGTTTTGGCTCCTGAATATTCCAAAGCCAAATCGGGGGCTGCGCAT 2886

Qy     926 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLys 945
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2887 CTGTTTGCCTTGTTGGAAAAGAAACCAAATATAGACAGCCGCAGTCAAGAAGGGAAAAAG 2946

Qy     946 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 965
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2947 CCAGACACATGTGAAGGGAATTTAGAGTTTCGAGAAGTCTCTTTCTTCTATCCATGTCGC 3006

Qy     966 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 985
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3007 CCAGATGTTTTCATCCTCCGTGGCTTATCCCTCAGTATTGAGCGAGGAAAGACAGTAGCA 3066

Qy     986 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr 1005
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3067 TTTGTGGGGAGCAGCGGCTGTGGGAAAAGCACTTCTGTTCAACTTCTGCAGAGACTTTAT 3126

Qy    1006 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 1025
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3127 GACCCCGTGCAAGGACAAGTGCTGTTTGATGGTGTGGATGCAAAAGAATTGAATGTACAG 3186

Qy    1026 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 1045
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    3187 TGGCTCCGTTCCCAAATAGCAATCGTTCCTCAAGAGCCTGTGCTCTTCAACTGCAGCATT 3246

Qy    1046 AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu 1065
```

```
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3247 GCTGAGAACATCGCCTATGGTGACAACAGCCGTGTGGTGCCATTAGATGAGATCAAAGAA 3306

Qy      1066 AlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThr 1085
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3307 GCCGCAAATGCAGCAAATATCCATTCTTTTATTGAAGGTCTCCCTGAGAAATACAACACA 3366

Qy      1086 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAla 1105
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3367 CAAGTTGGACTGAAAGGAGCACAGCTTTCTGGCGGCCAGAAACAAAGACTAGCTATTGCA 3426

Qy      1106 ArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsp 1125
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3427 AGGGCTCTTCTCCAAAAACCCAAAATTTTATTGTTGGATGAGGCCACTTCAGCCCTCGAT 3486

Qy      1126 AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys 1145
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3487 AATGACAGTGAGAAGGTGGTTCAGCATGCCCTTGATAAAGCCAGGACGGGAAGGACATGC 3546

Qy      1146 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 1165
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3547 CTAGTGGTCACTCACAGGCTCTCTGCAATTCAGAACGCAGATTTGATAGTGGTTCTGCAC 3606

Qy      1166 AsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyr 1185
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      3607 AATGGAAAGATAAAGGAACAAGGAACTCATCAAGAGCTCCTGAGAAATCGAGACATATAT 3666

Qy      1186 PheLysLeuValAsnAlaGlnSerValGln 1195
             ||||||||||||||||||||||||||||||
Db      3667 TTTAAGTTAGTGAATGCACAGTCAGTGCAG 3696
```

RESULT 2
AAD03506

```
ID   AAD03506 standard; cDNA; 4279 BP.
XX
AC   AAD03506;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) allelic variant (Genotype D) cDNA.
XX
KW   Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KW   MDR1; drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag= a
FT                   /product= "Dog P-glycoprotein (PGP) allelic variant
FT                   (Genotype D) protein"
FT   allele          replace (91, T)
FT                   /*tag= b
FT   allele          replace (607, C)
FT                   /*tag= c
FT   allele          replace (1001, T)
FT                   /*tag= c
FT   allele          replace (3458, A)
FT                   /*tag= c
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00310.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 9; Page 102-107; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) allelic variant
CC   (Genotype D) cDNA. The PGP enzyme functions as an efflux pump
CC   exporting small molecules across the cell membrane. This enzyme
CC   is a member of the ABC transporter family.
XX
SQ   Sequence 4279 BP; 1296 A; 833 C; 1009 G; 1141 T; 0 other;

Alignment Scores:
Pred. No.:              0               Length:         4279
Score:                  3557.00         Matches:        683
Percent Similarity:     74.74%          Conservative:   237
Best Local Similarity:  55.48%          Mismatches:     269
Query Match:            59.17%          Indels:         42
DB:                     22              Gaps:           8

US-09-873-409-6 (1-1195) x AAD03506 (1-4279)

Qy         1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
             |||:::::|||   :::|||:::::::::::|||||    |||||||||||   |||
Db       167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy        21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
             |||   |||   |||:::|||:::        :::      :::
Db       227 GTTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy        35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                                  :::           ||||||   |||
Db       287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy        45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                                  ||||||   ||||||||         |||:::
Db       347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy        60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
                |||||||||:::|||   |||   :::   |||   ||||||      :::||||||
Db       407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy        80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
             ||||||||||||||:::::::::   ||||:::||||||||||||||||      |||:::
Db       467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy       100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
             ||||||||||||||||||:::|||   |||:::   ||||||::::::::||||||||||||
Db       527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy       119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
             ||||||   :::   ||||||::::::::::||||||      |||      ||||||
Db       587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy       139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
                :::|||||||||||||||||||||   ||||:::   ||||||::::::      |||
Db       647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706
```

```
Qy    159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db    707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy    179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db    767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy    199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
Db    827 GGACAAAAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAGGAATTGGG 886

Qy    219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238
Db    887 ATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCA 946

Qy    239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
Db    947 TCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TAT 1000

Qy    259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
Db   1001 ACTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGA 1060

Qy    279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
Db   1061 CAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTC 1120

Qy    299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
Db   1121 AAGATAATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGAT 1180

Qy    319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
Db   1181 AATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAA 1240

Qy    339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
Db   1241 GTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTT 1300

Qy    359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
Db   1301 GGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCC 1360

Qy    379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
Db   1361 ACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTT 1420

Qy    399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
Db   1421 CGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAA 1480

Qy    419 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
Db   1481 AACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAA 1540

Qy    439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
Db   1541 GCCAATGCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAG 1600

Qy    459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
Db   1601 AGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTT 1660

Qy    479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
Db   1661 CGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAA 1720

Qy    499 SerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 518
Db   1721 GCAGTGGTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCT 1780

Qy    519 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 538
Db   1781 CATCGTTTGTCTACAGTTCGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATT 1840

Qy    539 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 558
Db   1841 GTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTC 1900

Qy    559 MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 578
Db   1901 ACAATGCAG---ACAAGAGGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAA 1957

Qy    579 ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--- 597
Db   1958 AGTGAAAGTGATGCCTTGGAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAA 2014

Qy    598 ----------------------------------------AspLysAlaGluGluSer 603
Db   2015 AGAAGATCAACTCGCAGGAGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACA 2074

Qy    604 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 623
Db   2075 AAAGAGGACTTGAATGAGAATGTACCTCCAGTTTCCTTCTGGAGGATTCTGAAGCTGAAC 2134

Qy    624 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 643
Db   2135 TCAACTGAATGGCCTTATTTTGTGGTTGGTATATTTTGTGCTATTATAAACGGAGGCCTG 2194

Qy    644 HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAsp 662
Db   2195 CAACCAGCATTTTCAATAATATTTTCAAGGATTATAGGGATCTTTACCCGAGATGAGGAT 2254

Qy    663 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 682
Db   2255 CCTGAAACAAAACGACAGAATAGTAACATGTTTTCTGTATTGTTTCTAGTCCTTGGAATT 2314

Qy    683 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 702
Db   2315 ATTTCTTTTATTACATTTTTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTC 2374

Qy    703 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 722
Db   2375 ACTAAGCGGCTTCGATACATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTT 2434

Qy    723 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 742
Db   2435 GATGACCCTAAAAACACCACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAA 2494

Qy    743 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 762
Db   2495 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGG 2554

Qy    763 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 782
Db   2555 ACAGGCATTATTATATCCTTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATT 2614

Qy    783 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 802
Db   2615 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 2674

Qy    803 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 822
Db   2675 AAAGATAAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTC 2734

Qy    823 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 842
Db   2735 CGAACTGTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTG 2794

Qy    843 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 862
Db   2795 CAAGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATC 2854

Qy    863 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 882
```

```
               :::    |||    :::||||||:::||||||        ||||||||||||||||||||:::
Db     2855 ACCCAGGCAATGATGTATTTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTG 2914

Qy      883 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 902
                           |||      :::    :::    :::||||||:::||||||    :::|||
Db     2915 GCAAATGAGTTCATGAACTTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGT 2974

Qy      903 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 922
            |||||||||:::|||:::                      ||||||:::|||:::|||||||||
Db     2975 GCCATGGCAGTGGGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCA 3034

Qy      923 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 942
            |||||||||:::          ::::::||||||    |||    |||||||||      |||
Db     3035 GCAGCCCACGTCATCATGATCATTGAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCAC 3094

Qy      943 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 962
            |||    ||||||:::|||    |||||||||:::    |||    ||||||    |||    |||
Db     3095 GGCCTCAAGCCAAATACGTTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTAT 3154

Qy      963 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 982
            |||    |||||||||:::    :::|||:::||||||||||||    :::::::::|||:::
Db     3155 CCCACTCGACCAGACATCCCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAG 3214

Qy      983 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 1002
            |||:::|||    ||||||||||||||||||||||||||||||    ||||||||||||:::
Db     3215 ACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAG 3274

Qy     1003 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 1022
            |||    ||||||:::    |||    ||||||    ||||||    :::    |||    |||
Db     3275 CGCTTCTATGACCCCTTGGCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTG 3334

Qy     1023 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 1042
            ||||||||||||||||||:::    :::    ||||||    |||||||||:::||||||:::
Db     3335 AATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGAC 3394

Qy     1043 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 1062
            ||||||||||||||||||||||||||||||||||||||||||||||||        :::|||
Db     3395 TGCAGCATTGCCGAGAACATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAG 3454

Qy     1063 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 1082
            |||    :::||||||          ||||||||||||    |||||||||    ||||||||||||
Db     3455 ATTGTGCAGGCAGCCAAGGAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAA 3514

Qy     1083 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1102
            |||||||||:::||||||    |||||||||||||||||||||||||||||||||:::
Db     3515 TACAACACCAGAGTAGGAGACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATT 3574

Qy     1103 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 1122
            ||||||||||||||||||:::::::::|||    ||||||||||||||||||||||||||||
Db     3575 GCCATAGCTCGCGCTCTTGTTAGACAGCCTCATATTTGCTTTTGGATGAAGCTACATCA 3634

Qy     1123 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1142
            ||||||||    :::||||||||||||||||||    ||||||||||||||||||    |||
Db     3635 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 3694

Qy     1143 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1162
            |||||||||:::|||:::    ||||||||||||    ||||||||||||||||||||||||
Db     3695 CGCACCTGCATTGTGATCGCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTG 3754

Qy     1163 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1182
            |||          |||||||||:::||||||    ||||||||||||:::||||||        :::
Db     3755 GTGTTTCAGAATGGCAAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGGCCCAGAAA 3814

Qy     1183 AspIleTyrPheLysLeuValAsnAlaGlnSer 1193
                ||||||||    :::|||:::    |||:::
Db     3815 GGCATCTATTTTTCCATGGTCAGTGTCCAGGCT 3847
```

RESULT 3
AAF86128
ID   AAF86128 standard; cDNA; 4195 BP.
XX
AC   AAF86128;
XX
DT   25-JUN-2001 (first entry)
XX
DE   Cynomologous monkey P-glycoprotein cDNA variant 1.
XX
KW   Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   efflux pump; ss.
XX
OS   Macaca fascicularis.
XX
FH   Key   Location/Qualifiers
FT   CDS   100..3951
FT   /*tag= a
FT   /product= "PGP"
FT   /note= "P-glycoprotein"
FT   misc_feature   376..384
FT   /*tag= b
FT   /note= "Insertion of 9 nucleotides relative to PGP
FT   allelic variant AAF86127"
XX
PN   WO200123565-A1.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26592.
XX
PR   28-SEP-1999;   99US-0156921.
PR   12-OCT-1999;   99US-0158818.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ, Steimel-Crespi DT, Crespi CL;
XX
DR   WPI; 2001-316136/33.
DR   P-PSDB; AAB81065.
XX
PT   Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein
PT   (PGP) and homologous PGP polypeptides are useful for predicting
PT   bioavailability of compound and increasing PGP transporter activity in
PT   cell -
XX
PS   Example 1; Page 59-65; 84pp; English.
XX
CC   This invention relates to a polynucleotide sequence encoding a
CC   cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the
CC   PGP protein. PGP, also known as multidrug transporter, MDR1 is a member
CC   of the ABC transporter superfamily. The enzyme serves as an efflux pump
CC   exporting small molecules across the cell membrane. The invention
CC   includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence
CC   and protein, and also that of an allelic variant. The PGP polynucleotide
CC   sequence is useful for increasing PGP transporter activity in a cell.
CC   Antisense sequences of the cDNA are useful for inhibiting PGP transport
CC   activity in a mammalian cell. They may also be used for increasing the
CC   bioavailability of a drug. The present sequence represents cDNA encoding
CC   cynomologous monkey P-glycoprotein. This sequence contains a 9 nucleotide
CC   insert compared to the PGP allelic variant given in AAF86127.
XX
SQ   Sequence 4195 BP; 1230 A; 801 C; 1039 G; 1125 T; 0 other;

Alignment Scores:			
Pred. No.:	0	Length:	4195
Score:	3556.50	Matches:	680
Percent Similarity:	75.16%	Conservative:	246
Best Local Similarity:	55.19%	Mismatches:	263
Query Match:	59.16%	Indels:	43
DB:	22	Gaps:	9

US-09-873-409-6 (1-1195) x AAF86128 (1-4195)

```
Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            |||::::::|||    ||||||:::::::::::::||||||    ||||||||||||    |||
Db      250 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGAGCTGGACTTCCTCTCATGATGCTG 309

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIleSer---------------------GlyCys 33
```

```
Db     310 GTGTTTGGAGACATGACGGATACCTTTGCAAATGCAGGAAATTTAGGAGATTTAGGAGCT 369

Qy      34 LeuValGlnThrAsnThrTyrSer------------------------------------ 41

Db     370 CTGTTGTTTAACAACACTAATAGCAGTAATATCACTGATACAGTGCCCGTCATGAATCTG 429

Qy      42 ---------PhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIle 58

Db     430 GAGGAAGATATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTT 489

Qy      59 PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIle 78

Db     490 GCTGCTTACATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATT 549

Qy      79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp 98

Db     550 AGAAAACAGTTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGAT 609

Qy      99 IleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly 117

Db     610 GTTGGGGAGCTTAACACCCGGCTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGT 669

Qy     118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137

Db     670 GACAAAATTGGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGA 729

Qy     138 LeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157

Db     730 TTTACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTG 789

Qy     158 SerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr 177

Db     790 TCAGCTGCAGTCTGGGCAAAGATACTGTCTTCATTTACTGATAAAGAACTCTTAGCTTAT 849

Qy     178 SerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197

Db     850 GCAAAAGCTGGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTT 909

Qy     198 ArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 217

Db     910 GGAGGACAAAAGAAAGAACTCGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATT 969

Qy     218 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 237

Db     970 GGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTTATCTAT 1029

Qy     238 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 257

Db    1030 GCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAAAGGAA------ 1083

Qy     258 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 277

Db    1084 TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTT 1143

Qy     278 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 297

Db    1144 GGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTTTGAAATC 1203

Qy     298 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 317

Db    1204 TTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCA 1263

Qy     318 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 337

Db    1264 GATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAA 1323

Qy     338 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 357

Db    1324 GAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTG 1383

Qy     358 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 377

Db    1384 GTTGGAAACAGCGGCTGTGGGAAGAGCACAACGGTCCAGCTGATGCAGAGGCTTTATGAC 1443

Qy     378 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 397

Db    1444 CCCACAGAGGGCATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAACGTAAGGTTT 1503

Qy     398 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 417

Db    1504 CTACGGGAAATCATCGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCT 1563

Qy     418 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 437

Db    1564 GAAAACATTCGCTATGGTCGTGAAGATGTCACCATGGATGAGATTGAGAAAGCTGTCAAG 1623

Qy     438 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 457

Db    1624 GAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCAGAAATTTGACACCCTGGTTGGA 1683

Qy     458 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 477

Db    1684 GAGAGAGGGGCCCAGCTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTG 1743

Qy     478 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 497

Db    1744 GTTCGCAACCCCAAGATCCTCCTGCTGGACGAGGCCACGTCAGCCTTGGACACAGAAAGT 1803

Qy     498 LysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValVal 517

Db    1804 GAAGCAGTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATA 1863

Qy     518 AlaHisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMet 537

Db    1864 GCTCATCGTTTGTCTACGGTTCGTAATGCCGACGTCATCGCTGGTTTCGATGATGGAGTC 1923

Qy     538 LeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeu 557

Db    1924 ATTGTGGAGAAAGGAAATCATGATGAGCTCATGAAAGAGAAAGGCATTTACTTCAAACTT 1983

Qy     558 ValMetSerGlnAspIle--------------------LysLysAlaAspGlu------ 568

Db    1984 GTCACAATGCAGACAGCAGGAAATGAAATTGAATTAGAAAATGCAGCTGATGAATCCAAA 2043

Qy     569 ---GlnMetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu--------- 584

Db    2044 AGTGAAATTGATACCTTGGAAATGTCTTCACATGATTCAGGATCCAGTCTAATAAGAAAA 2103

Qy     585 -----ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGlu 602

Db    2104 AGATCCACTCGTAGGAGTGTCCGTGGATCACAAGGCCAA------GACAGAAAGCTTAGT 2157

Qy     603 SerThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeu 622

Db    2158 ACCAAAGAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTA 2217

Qy     623 AsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThr 642

Db    2218 AATTTAACTGAGTGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGT 2277

Qy     643 ValHisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsn 661

Db    2278 CTGCAACCAGCATTTGCAGTAATATTTTCAAAGATTATAGGGATTTTTACAAGAAATGAT 2337

Qy     662 AspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGly 681

Db    2338 GATGCCGAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGTCCTTGGA 2397

Qy     682 ValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIle 701

Db    2398 ATTGTTTCTTTTATTACATTTTTCCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATC 2457

Qy     702 LeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrp 721

Db    2458 CTCACCAAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGCTGG 2517
```

```
Qy    722 PheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAla 741
Db   2518 TTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCT 2577

Qy    742 GlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMet 761
Db   2578 CAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTATAATTACCCAGAATATAGCAAATCTT 2637

Qy    762 GlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSer 781
Db   2638 GGGACAGGAATAATTATATCCTTAATCTATGGTTGGCAACTGACACTGTTACTCTTAGCA 2697

Qy    782 IleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAla 801
Db   2698 ATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCA 2757

Qy    802 AsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsn 821
Db   2758 CTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAAC 2817

Qy    822 IleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMet 841
Db   2818 TTCCGAACTGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGATCAGAGT 2877

Qy    842 LeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAla 861
Db   2878 TTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATCACGTTTTCC 2937

Qy    862 PheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeu 881
Db   2938 TTCACGCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTG 2997

Qy    882 IleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyr 901
Db   2998 GTGGCACATAGTCTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTT 3057

Qy    902 GlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSer 921
Db   3058 GGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTA 3117

Qy    922 GlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGln 941
Db   3118 TCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACA 3177

Qy    942 GluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePhe 961
Db   3178 GAAGGCCTAAAGCCGAACACATTGGAAGGAAATGTCACATTTAATGAAGTTGTATTCAAC 3237

Qy    962 TyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGly 981
Db   3238 TATCCCACCCGACTGGACATCCCAGTGCTTCAGGGGCTGAGCCTGGAAGTGAAGAAGGGC 3297

Qy    982 LysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeu 1001
Db   3298 CAGACGCTGGCCCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACGGTGGTCCAGCTCCTG 3357

Qy   1002 GlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGlu 1021
Db   3358 GAGCGGTTCTATGACCCCTTGGCGGGGAAAGTGCTGCTTGACGGCAAAGAAATAAAGCAA 3417

Qy   1022 LeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPhe 1041
Db   3418 CTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTT 3477

Qy   1042 AsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAsp 1061
Db   3478 GACTGCAGCATTAGTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAA 3537

Qy   1062 GluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGlu 1081
Db   3538 GAGATCGTGAGGGCAGCCAAGGAGGCCAATATACACGCCTTCATCGAGTCACTGCCTAAT 3597

Qy   1082 LysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArg 1101
Db   3598 AAATATAGCACCAGAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGC 3657

Qy   1102 LeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThr 1121
Db   3658 ATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACA 3717

Qy   1122 SerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThr 1141
Db   3718 TCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAA 3777

Qy   1142 GlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIle 1161
Db   3778 GGCCGTACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATA 3837

Qy   1162 ValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsn 1181
Db   3838 GTGGTGTTTCAGAATGGCAGAGTCAAGGAGCACGGCACACATCAGCAGCTGCTGGCACAG 3897

Qy   1182 ArgAspIleTyrPheLysLeuValAsnAlaGlnSer 1193
Db   3898 AAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGCT 3933
```

```
RESULT 4
AAZ49332
ID   AAZ49332 standard; cDNA; 3860 BP.
XX
AC   AAZ49332;
XX
DT   14-MAR-2000  (first entry)
XX
DE   Human wild-type multidrug resistance-1 (MDR-1) cDNA.
XX
KW   Multidrug resistance; MDR-1; P-glycoprotein;
KW   transmembrane efflux pump; haematopoietic stem cell; transduction;
KW   bone marrow transplantation; chemotherapy; radiation therapy; cancer;
KW   gene therapy; gene replacement; genetic defect; thalassaemia;
KW   Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
KW   cytokine; wild-type; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3843
FT                   /*tag=  a
FT                   /product= "Human wild-type MDR-1 protein"
FT   mutation        replace (553..555, GTT)
FT                   /*tag=  b
FT                   /note= "cDNA sequence of G185V human mutant MDR-1 given
FT                          in AAZ49333"
XX
PN   WO9961589-A2.
XX
PD   02-DEC-1999.
XX
PF   27-MAY-1999;   99WO-US11825.
XX
PR   28-MAY-1998;   98US-0086988.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI   Sorrentino B,  Bunting K;
XX
DR   WPI; 2000-072615/06.
DR   P-PSDB; AAY58186.
XX
PT   Ex vivo expansion of hematopoietic stem cells transduced with a
PT   sequence encoding human multidrug resistance-1, used for bone marrow
PT   transplantation -
XX
PS   Claim 10; Page 68-70; 113pp; English.
XX
```

```
CC   This sequence represents cDNA encoding human wild-type
CC   multidrug resistance protein MDR-1. MDR-1 is a transmembrane
CC   efflux pump, responsible for the export of drugs from cells,
CC   particularly cancer cells. Wild-type MDR-1 shows increased
CC   resistance to etoposide and decreased resistance to vinca
CC   alkaloids compared with a mutant form (AAY58187) where the Gly at
CC   position 185 is replaced by Val. The invention relates to transducing
CC   haematopoietic stem cells with nucleic acid encoding an MDR protein
CC   and culturing the modified cells. The modified haematopoietic stem
CC   cells are useful in bone marrow transplantation (to reconstitute
CC   haematopoietic systems in patients who have undergone chemotherapy or
CC   radiation therapy) and in ex vivo gene therapy of genetic defects in
CC   cells derived from haematopoietic stem cells, e.g., thalassaemia,
CC   Gauchier's disease, sickle cell anaemia or leukaemia. The modified
CC   cells can also be used to identify factors involved in regulating
CC   proliferation and differentiation in haematopoietic stem cells.
CC   Haematopoietic stem cells that express MDR-1 will be protected against
CC   chemotherapeutic agents, so can be engrafted while the patient is
CC   undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells
CC   provides sufficient cells to permit standard biochemical analysis.
CC   Overexpression of MDR-1 allows cytokine-driven expansion of
CC   haematopoietic stem cells by at least 10-fold compared with a maximum
CC   of 4-fold in known procedures.
XX
SQ   Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;

Alignment Scores:
Pred. No.:               0          Length:        3860
Score:                   3554.00    Matches:       682
Percent Similarity:      75.63%     Conservative:  246
Best Local Similarity:   55.58%     Mismatches:    263
Query Match:             59.12%     Indels:        36
DB:                      21         Gaps:          9

US-09-873-409-6 (1-1195) x AAZ49332 (1-3860)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    151 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210

Qy     21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
Db    211 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 270

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db    271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    331 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 390

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    451 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 510

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    511 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGGTATTGGTGACAAAATT 570

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 630

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    631 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    691 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db    811 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 870

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db    871 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 930

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db    931 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 984

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db    985 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1044

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db   1045 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db   1105 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1164

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db   1165 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1224

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   1225 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1284

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   1285 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1344

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   1345 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1404

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   1405 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1464

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1465 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1524

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1525 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1584

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   1585 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1644

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   1645 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1704

Qy    501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db   1705 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 1764

Qy    521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
```

```
Db   1765 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 1824

Qy    541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db   1825 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 1884

Qy    561 Gln------------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
Db   1885 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 1944

Qy    575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 589
Db   1945 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2004

Qy    590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
Db   2005 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2064

Qy    608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627
Db   2065 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2124

Qy    628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
Db   2125 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2184

Qy    648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
Db   2185 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2244

Qy    667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
Db   2245 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2304

Qy    687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
Db   2305 ACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2364

Qy    707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
Db   2365 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2424

Qy    727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
Db   2425 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2484

Qy    747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
Db   2485 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2544

Qy    767 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
Db   2545 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 2604

Qy    787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
Db   2605 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 2664

Qy    807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
Db   2665 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 2724

Qy    827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
Db   2725 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2784

Qy    847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
Db   2785 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 2844

Qy    867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
Db   2845 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 2904

Qy    887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
Db   2905 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 2964

Qy    907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
Db   2965 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3024

Qy    927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
Db   3025 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3084

Qy    947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
Db   3085 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3144

Qy    967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
Db   3145 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3204

Qy    987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
Db   3205 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3264

Qy   1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
Db   3265 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3324

Qy   1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
Db   3325 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3384

Qy   1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
Db   3385 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3444

Qy   1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
Db   3445 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3504

Qy   1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
Db   3505 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3564

Qy   1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
Db   3565 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 3624

Qy   1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
Db   3625 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 3684

Qy   1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
Db   3685 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 3744

Qy   1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
Db   3745 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 3804

Qy   1187 LysLeuValAsnAlaGlnSer 1193
Db   3805 TCAATGGTCAGTGTCCAGGCT 3825
```

RESULT 5
ABA94365
ID   ABA94365 standard; DNA; 3860 BP.
XX
AC   ABA94365;
XX
DT   26-MAR-2002 (first entry)
XX

```
DE   Human BCRP DNA related seq Id No. 1.
XX
KW   Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;
KW   hepatotropic; neurodegenerative; cytostatic; antianemic; muscular; BCRP;
KW   cardiant; gene therapy; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3843
FT                   /*tag=  a
XX
PN   WO200192877-A2.
XX
PD   06-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US17459.
XX
PR   31-MAY-2000; 2000US-0584586.
PR   29-MAY-2001; 2001US-0866866.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI   Sorrentino B,  Schuetz J;
XX
DR   WPI; 2002-114368/15.
DR   P-PSDB; ABB07266.
XX
PT   Identifying a stem cell, for treating e.g., muscular dystrophy,
PT   myocardial infarction, Parkinson's disease, or neurodegenerative
PT   disorders, comprises detecting the expression of an ATP transport
PT   protein (BCRP) by a cell  -
XX
PS   Disclosure; Page 53-55; 87pp; English.
XX
CC   The invention provides a method of identifying and/or isolating a stem
CC   cell that involves detecting the expression of an ATP transport protein
CC   containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
CC   comprising stem cells. The isolated stem cells may be used in the
CC   treatment of diseases such as muscular dystrophy, degenerative liver
CC   disorder, myocardial infarction, Parkinson's disease, degenerative
CC   disorders of the brain, and for tissue regeneration or replacement.
CC   Haematopoietic cells can be used in bone marrow transplants (e.g., for
CC   treatment of leukemia) and for ex vivo gene therapy for treating blood
CC   diseases such as sickle cell anemia and thalassemia. The stem cells can
CC   also be used as cell targets in gene therapy protocols. The present
CC   sequence represents a sequence related to the BCRP for which no relevant
CC   information has been provided in the specification.
XX
SQ   Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;

Alignment Scores:
Pred. No.:              0             Length:        3860
Score:                  3554.00       Matches:       682
Percent Similarity:     75.63%        Conservative:  246
Best Local Similarity:  55.58%        Mismatches:    263
Query Match:            59.12%        Indels:        36
DB:                     24            Gaps:          9

US-09-873-409-6 (1-1195) x ABA94365 (1-3860)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    151 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210

Qy     21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
Db    211 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 270

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db    271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    331 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 390

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    451 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 510

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    511 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGGTATTGGTGACAAAATT 570

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 630

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    631 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    691 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db    811 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 870

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db    871 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 930

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db    931 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 984

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db    985 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1044

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db   1045 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db   1105 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1164

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db   1165 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1224

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   1225 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1284

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   1285 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1344

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   1345 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1404

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
```

```
Db   1405 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1464

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1465 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1524

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1525 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1584

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   1585 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1644

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   1645 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1704

Qy    501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db   1705 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 1764

Qy    521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db   1765 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 1824

Qy    541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db   1825 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 1884

Qy    561 Gln------------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
Db   1885 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 1944

Qy    575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 589
Db   1945 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2004

Qy    590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
Db   2005 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2064

Qy    608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627
Db   2065 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2124

Qy    628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
Db   2125 CCTTATTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2184

Qy    648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
Db   2185 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2244

Qy    667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
Db   2245 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2304

Qy    687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
Db   2305 ACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2364

Qy    707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
Db   2365 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2424

Qy    727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
Db   2425 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2484

Qy    747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
Db   2485 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2544

Qy    767 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
Db   2545 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 2604

Qy    787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
Db   2605 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 2664

Qy    807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
Db   2665 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 2724

Qy    827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
Db   2725 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2784

Qy    847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
Db   2785 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 2844

Qy    867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
Db   2845 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 2904

Qy    887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
Db   2905 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 2964

Qy    907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
Db   2965 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3024

Qy    927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
Db   3025 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3084

Qy    947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
Db   3085 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3144

Qy    967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
Db   3145 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3204

Qy    987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
Db   3205 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3264

Qy   1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
Db   3265 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3324

Qy   1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
Db   3325 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3384

Qy   1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
Db   3385 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3444

Qy   1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
Db   3445 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3504

Qy   1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
Db   3505 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3564

Qy   1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
Db   3565 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 3624
```

```
Qy     1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
              :::|||||||||||||||||   ||||||||||||||||    |||||||||||:::
Db     3625 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 3684

Qy     1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
              |||:::   ||||||||||   |||||||||||||||||||||||||         |||
Db     3685 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 3744

Qy     1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
              |||::::::|||||   |||||||||||:::||||||       :::   ||||||||
Db     3745 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 3804

Qy     1187 LysLeuValAsnAlaGlnSer 1193
                 :::|||:::   |||:::
Db     3805 TCAATGGTCAGTGTCCAGGCT 3825

RESULT 6
AAD03504
ID   AAD03504 standard; cDNA; 4279 BP.
XX
AC   AAD03504;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) allelic variant (Genotype A) cDNA.
XX
KW   Dog; P-glycoprotein allelic variant; multidrug transporter; MDR1;
KW   drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag=  a
FT                   /product= "Dog P-glycoprotein (PGP) allelic variant
FT                   (Genotype A) protein"
FT   allele          replace (607, C)
FT                   /*tag=  b
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;    99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00308.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 9; Page 85-90; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP .
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) allelic variant
CC   (Genotype A) cDNA. The PGP enzyme functions as an efflux pump
CC   exporting small molecules across the cell membrane. This enzyme
CC   is a member of the ABC transporter family.
XX
SQ   Sequence 4279 BP; 1295 A; 833 C; 1008 G; 1143 T; 0 other;

Alignment Scores:
Pred. No.:              0              Length:        4279
Score:                  3554.00        Matches:       682
Percent Similarity:     74.74%         Conservative:  238
Best Local Similarity:  55.40%         Mismatches:    269
Query Match:            59.12%         Indels:        42
DB:                     22             Gaps:          8

US-09-873-409-6 (1-1195) x AAD03504 (1-4279)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            |||::::::|||   :::|||::::::::::::||||||   |||||||||||   |||
Db      167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
            |||   |||   |||:::|||:::         :::      :::
Db      227 GTTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy       35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                                 :::      ||||||       |||
Db      287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy       45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                                 ||||||    |||||||||          |||:::
Db      347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy       60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
               |||||||||:::|||   |||   :::   |||   ||||||       :::||||||
Db      407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy       80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
            |||||||||||||||:::::::::   |||:::|||||||||||||||       |||:::
Db      467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy      100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
            ||||||||||||||||||:::|||   |||:::   ||||||::::::||||||||||||
Db      527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy      119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
            ||||||   :::   ||||||:::::::::||||||      |||      ||||||
Db      587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy      139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
               :::||||||||||||||||||||   |||:::   ||||||::::::      |||
Db      647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy      159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
            ||||||      ::::::::::::   |||   |||   |||||||||   ||||||:::
Db      707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy      179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
            |||||||||||||||||||||||||||||||::::::||||||||||||||||||||||||
Db      767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy      199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
            |||:::||||||||||:::||||||   :::||||||::::::||||||         |||
Db      827 GGACAAAAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAGGAATTGGG 886

Qy      219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238
            ||||||:::   |||   :::   :::|||:::||||||   :::         :::
Db      887 ATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCA 946

Qy      239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
            :::|||   |||||||||||||||||||||||||||:::|||:::   |||      |||
Db      947 TCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TAT 1000
```

```
Qy    259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
Db   1001 TCTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGA 1060

Qy    279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
Db   1061 CAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTC 1120

Qy    299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
Db   1121 AAGATAATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGAT 1180

Qy    319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
Db   1181 AATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAA 1240

Qy    339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
Db   1241 GTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTT 1300

Qy    359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
Db   1301 GGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCC 1360

Qy    379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
Db   1361 ACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTT 1420

Qy    399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
Db   1421 CGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAA 1480

Qy    419 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
Db   1481 AACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAA 1540

Qy    439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
Db   1541 GCCAATGCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAG 1600

Qy    459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
Db   1601 AGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTT 1660

Qy    479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
Db   1661 CGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAA 1720

Qy    499 SerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 518
Db   1721 GCAGTGGTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCT 1780

Qy    519 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 538
Db   1781 CATCGTTTGTCTACAGTTCGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATT 1840

Qy    539 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 558
Db   1841 GTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTC 1900

Qy    559 MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 578
Db   1901 ACAATGCAG---ACAAGAGGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAA 1957

Qy    579 ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--- 597
Db   1958 AGTGAAAGTGATGCCTTGGAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAA 2014

Qy    598 ------------------------------------------AspLysAlaGluGluSer 603
Db   2015 AGAAGATCAACTCGCAGGAGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACA 2074

Qy    604 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 623
Db   2075 AAAGAGGACTTGAATGAGAATGTACCTCCAGTTTCCTTCTGGAGGATTCTGAAGCTGAAC 2134

Qy    624 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 643
Db   2135 TCAACTGAATGGCCTTATTTTGTGGTTGGTATATTTTGTGCTATTATAAACGGAGGCCTG 2194

Qy    644 HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAsp 662
Db   2195 CAACCAGCATTTTCAATAATATTTTCAAGGATTATAGGGATCTTTACCCGAGATGAGGAT 2254

Qy    663 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 682
Db   2255 CCTGAAACAAAACGACAGAATAGTAACATGTTTTCTGTATTGTTTCTAGTCCTTGGAATT 2314

Qy    683 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 702
Db   2315 ATTTCTTTTATTACATTTTTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTC 2374

Qy    703 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 722
Db   2375 ACTAAGCGGCTTCGATACATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTT 2434

Qy    723 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 742
Db   2435 GATGACCCTAAAAACACCACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAA 2494

Qy    743 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 762
Db   2495 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGG 2554

Qy    763 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 782
Db   2555 ACAGGCATTATTATATCCTTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATT 2614

Qy    783 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 802
Db   2615 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 2674

Qy    803 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 822
Db   2675 AAAGATAAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTC 2734

Qy    823 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 842
Db   2735 CGAACTGTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTG 2794

Qy    843 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 862
Db   2795 CAAGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATC 2854

Qy    863 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 882
Db   2855 ACCCAGGCAATGATGTATTTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTG 2914

Qy    883 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 902
Db   2915 GCAAATGAGTTCATGAACTTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGT 2974

Qy    903 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 922
Db   2975 GCCATGGCAGTGGGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCA 3034

Qy    923 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 942
Db   3035 GCAGCCCACGTCATCATGATCATTGAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCAC 3094

Qy    943 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 962
Db   3095 GGCCTCAAGCCAAATACGTTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTAT 3154

Qy    963 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 982
```

```
            |||   |||||||||:::    :::|||:::|||||||||||    :::::::::|||:::
Db     3155 CCCACTCGACCAGACATCCCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAG 3214

Qy      983 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 1002
            |||:::|||   ||||||||||||||||||||||||||||||     ||||||||||:::
Db     3215 ACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAG 3274

Qy     1003 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 1022
            |||   |||||||||:::   |||   ||||||   ||||||   :::   |||   |||
Db     3275 CGCTTCTATGACCCCTTGGCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTG 3334

Qy     1023 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 1042
            ||||||||||||||||||:::    :::   ||||||   |||||||||:::||||||:::
Db     3335 AATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGAC 3394

Qy     1043 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 1062
            ||||||||||||||||||||||||||||||||||||||||||||||||      :::|||
Db     3395 TGCAGCATTGCCGAGAACATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAG 3454

Qy     1063 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 1082
            |||   :::||||||      ||||||||||||   |||||||||   ||||||||||||
Db     3455 ATTATGCAGGCAGCCAAGGAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAA 3514

Qy     1083 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1102
            |||||||||:::||||||      ||||||   ||||||||||||||||||||||||:::
Db     3515 TACAACACCAGAGTAGGAGACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATT 3574

Qy     1103 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 1122
            ||||||||||||||||||:::::::::|||   ||||||||||||||||||||||||
Db     3575 GCCATAGCTCGCGCTCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCTACATCA 3634

Qy     1123 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1142
            |||||||||   :::||||||||||||||||||   ||||||||||||||||||   |||
Db     3635 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 3694

Qy     1143 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1162
            |||||||||:::|||:::   ||||||||||||   ||||||||||||||||||||||||
Db     3695 CGCACCTGCATTGTGATCGCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTG 3754

Qy     1163 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1182
            |||      |||||||||:::||||||   ||||||||||||:::||||||      :::
Db     3755 GTGTTTCAGAATGGCAAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGGCCCAGAAA 3814

Qy     1183 AspIleTyrPheLysLeuValAsnAlaGlnSer 1193
               |||||||||   :::|||:::   |||:::
Db     3815 GGCATCTATTTTTCCATGGTCAGTGTCCAGGCT 3847
```

RESULT 7
AAD03505

```
ID   AAD03505 standard; cDNA; 4279 BP.
XX
AC   AAD03505;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.
XX
KW   Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KW   MDR1; drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag= a
FT                   /product= "Dog P-glycoprotein (PGP) allelic variant
FT                   (Genotype B) protein"
FT   allele          replace (91, T)
FT                   /*tag= b
FT   allele          replace (607, C)
FT                   /*tag= c
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00309.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 9; Page 93-99; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) allelic variant
CC   (Genotype B) cDNA. The PGP enzyme functions as an efflux pump
CC   exporting small molecules across the cell membrane. This enzyme
CC   is a member of the ABC transporter family.
XX
SQ   Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;
```

```
Alignment Scores:
Pred. No.:              0             Length:        4279
Score:                  3554.00       Matches:       682
Percent Similarity:     74.74%        Conservative:  238
Best Local Similarity:  55.40%        Mismatches:    269
Query Match:            59.12%        Indels:        42
DB:                     22            Gaps:          8
```

US-09-873-409-6 (1-1195) x AAD03505 (1-4279)

```
Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            |||::::::|||    :::|||::::::::::::||||||   ||||||||||||   |||
Db      167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
            |||   |||   |||:::|||:::        :::      :::
Db      227 GTTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy       35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                                 :::      ||||||      |||
Db      287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy       45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                           ||||||   |||||||||         |||:::
Db      347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy       60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
               |||||||||:::|||   |||   :::   |||   ||||||      :::||||||
Db      407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy       80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
            ||||||||||||||||||::::::::   |||:::|||||||||||||||      |||:::
Db      467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526
```

```
Qy    100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db    527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy    119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db    587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy    139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db    647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy    159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db    707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy    179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db    767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy    199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
Db    827 GGACAAAAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAGGAATTGGG 886

Qy    219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238
Db    887 ATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCA 946

Qy    239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
Db    947 TCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TAT 1000

Qy    259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
Db   1001 TCTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGA 1060

Qy    279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
Db   1061 CAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTC 1120

Qy    299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
Db   1121 AAGATAATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGAT 1180

Qy    319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
Db   1181 AATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAA 1240

Qy    339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
Db   1241 GTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTT 1300

Qy    359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
Db   1301 GGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCC 1360

Qy    379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
Db   1361 ACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTT 1420

Qy    399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
Db   1421 CGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAA 1480

Qy    419 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
Db   1481 AACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAA 1540

Qy    439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
Db   1541 GCCAATGCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAG 1600

Qy    459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
Db   1601 AGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTT 1660

Qy    479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
Db   1661 CGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAA 1720

Qy    499 SerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 518
Db   1721 GCAGTGGTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCT 1780

Qy    519 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 538
Db   1781 CATCGTTTGTCTACAGTTCGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATT 1840

Qy    539 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 558
Db   1841 GTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTC 1900

Qy    559 MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 578
Db   1901 ACAATGCAG---ACAAGAGGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAA 1957

Qy    579 ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--- 597
Db   1958 AGTGAAAGTGATGCCTTGGAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAA 2014

Qy    598 ------------------------------------------AspLysAlaGluGluSer 603
Db   2015 AGAAGATCAACTCGCAGGAGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACA 2074

Qy    604 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 623
Db   2075 AAAGAGGACTTGAATGAGAATGTACCTCCAGTTTCCTTCTGGAGGATTCTGAAGCTGAAC 2134

Qy    624 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 643
Db   2135 TCAACTGAATGGCCTTATTTTGTGGTTGGTATATTTTGTGCTATTATAAACGGAGGCCTG 2194

Qy    644 HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAsp 662
Db   2195 CAACCAGCATTTTCAATAATATTTTCAAGGATTATAGGATCTTTACCCGAGATGAGGAT 2254

Qy    663 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 682
Db   2255 CCTGAAACAAAACGACAGAATAGTAACATGTTTTCTGTATTGTTTCTAGTCCTTGGAATT 2314

Qy    683 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 702
Db   2315 ATTTCTTTTATTACATTTTTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTC 2374

Qy    703 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 722
Db   2375 ACTAAGCGGCTTCGATACATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTT 2434

Qy    723 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 742
Db   2435 GATGACCCTAAAAACACCACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAA 2494

Qy    743 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 762
Db   2495 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGG 2554

Qy    763 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 782
Db   2555 ACAGGCATTATTATATCCTTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATT 2614

Qy    783 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 802
Db   2615 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 2674

Qy    803 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 822
```

```
            ||||||||:::||||||:::    ||||||||||||||||||||||:::||||||
Db     2675 AAAGATAAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTC 2734

Qy      823 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 842
            ||||||:::||||||||||||||||||:::   ||||||    ||||||   :::   |||
Db     2735 CGAACTGTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTG 2794

Qy      843 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 862
            |||      :::||||||:::   :::||||||   |||   |||      ::::::
Db     2795 CAAGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATC 2854

Qy      863 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 882
            :::   |||   :::||||||:::||||||         ||||||||||||||||||:::
Db     2855 ACCCAGGCAATGATGTATTTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTG 2914

Qy      883 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 902
                     |||      :::   :::   :::||||||:::||||||   :::|||
Db     2915 GCAAATGAGTTCATGAACTTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGT 2974

Qy      903 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 922
            ||||||||:::||||:::                ||||||:::|||:::|||||||||
Db     2975 GCCATGGCAGTGGGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCA 3034

Qy      923 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 942
            |||||||||:::      ::::::||||||   |||   |||||||||   |||
Db     3035 GCAGCCCACGTCATCATGATCATTGAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCAC 3094

Qy      943 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 962
            |||   ||||||:::|||   |||||||||:::   |||   ||||||   |||   |||
Db     3095 GGCCTCAAGCCAAATACGTTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTAT 3154

Qy      963 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 982
            |||   |||||||||:::   :::|||:::||||||||||||   :::::::::|||:::
Db     3155 CCCACTCGACCAGACATCCCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAG 3214

Qy      983 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 1002
            |||:::|||   |||||||||||||||||||||||||||||||   ||||||||||||:::
Db     3215 ACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAG 3274

Qy     1003 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 1022
            |||   |||||||||:::   ||||||   ||||||   |||   :::   |||   |||
Db     3275 CGCTTCTATGACCCCTTGGCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTG 3334

Qy     1023 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 1042
            ||||||||||||||||||:::   :::   ||||||   |||||||||:::||||||:::
Db     3335 AATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGAC 3394

Qy     1043 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 1062
            ||||||||||||||||||||||||||||||||||||||||||||||||      :::|||
Db     3395 TGCAGCATTGCCGAGAACATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAG 3454

Qy     1063 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 1082
            |||   :::||||||      ||||||||||||   |||||||||   ||||||||||||
Db     3455 ATTATGCAGGCAGCCAAGGAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAA 3514

Qy     1083 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1102
            |||||||||:::||||||   ||||||   ||||||||||||||||||||||||||||:::
Db     3515 TACAACACCAGAGTAGGAGACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATT 3574

Qy     1103 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 1122
            |||||||||||||||||||||:::::::|||   ||||||||||||||||||||||||
Db     3575 GCCATAGCTCGCGCTCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCTACATCA 3634

Qy     1123 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1142
            |||||||||   :::||||||||||||||||||   ||||||||||||||||||   |||
Db     3635 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 3694

Qy     1143 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1162
            |||||||||:::||||:::   ||||||||||||   ||||||||||||||||||||||||
Db     3695 CGCACCTGCATTGTGATCGCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTG 3754

Qy     1163 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1182
            |||      |||||||||:::||||||   ||||||||||||:::||||||      :::
Db     3755 GTGTTTCAGAATGGCAAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGGCCCAGAAA 3814

Qy     1183 AspIleTyrPheLysLeuValAsnAlaGlnSer 1193
            |||||||||   :::|||:::   |||:::
Db     3815 GGCATCTATTTTTCCATGGTCAGTGTCCAGGCT 3847
```

RESULT 8
AAD03488

```
ID   AAD03488 standard; cDNA; 4279 BP.
XX
AC   AAD03488;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) cDNA #1.
XX
KW   Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag=  a
FT                   /product= "Dog P-glycoprotein (PGP) #1"
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00303.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 3; Page 58-63; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) cDNA. This
CC   sequence is also referred as Genotype C cDNA. The
CC   PGP enzyme functions as an efflux pump exporting small molecules
CC   across the cell membrane. This enzyme is a member of the ABC
CC   transporter family.
XX
SQ   Sequence 4279 BP; 1294 A; 834 C; 1008 G; 1143 T; 0 other;
```

Alignment Scores:

Pred. No.:	0	Length:	4279
Score:	3549.00	Matches:	681
Percent Similarity:	74.65%	Conservative:	238
Best Local Similarity:	55.32%	Mismatches:	270

```
Query Match:            59.03%     Indels:        42
DB:                     22         Gaps:          8

US-09-873-409-6 (1-1195) x AAD03488 (1-4279)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
Db    227 GTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy     35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
Db    287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy     45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db    347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy     60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db    407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy     80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db    467 AAACAATTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy    100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db    527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy    119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db    587 AAAATTGGAATGTTCTTTCACTCAATAGCAACATTTTTCACCGGTTTATAGTGGGGTTT 646

Qy    139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db    647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy    159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db    707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy    179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db    767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy    199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
Db    827 GGACAAAAGAAAGAACTTGAAAGGTACAACAAAAATTAGAAGAAGCTAAAGGAATTGGG 886

Qy    219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238
Db    887 ATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCA 946

Qy    239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
Db    947 TCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TAT 1000

Qy    259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
Db   1001 TCTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGA 1060

Qy    279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
Db   1061 CAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTC 1120

Qy    299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
Db   1121 AAGATAATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGAT 1180

Qy    319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
Db   1181 AATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAA 1240

Qy    339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
Db   1241 GTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTT 1300

Qy    359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
Db   1301 GGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCC 1360

Qy    379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
Db   1361 ACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTT 1420

Qy    399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
Db   1421 CGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAA 1480

Qy    419 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
Db   1481 AACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAA 1540

Qy    439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
Db   1541 GCCAATGCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAG 1600

Qy    459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
Db   1601 AGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTT 1660

Qy    479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
Db   1661 CGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAA 1720

Qy    499 SerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 518
Db   1721 GCAGTGGTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCT 1780

Qy    519 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 538
Db   1781 CATCGTTTGTCTACAGTTCGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATT 1840

Qy    539 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 558
Db   1841 GTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTC 1900

Qy    559 MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 578
Db   1901 ACAATGCAG---ACAAGAGGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAA 1957

Qy    579 ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--- 597
Db   1958 AGTGAAAGTGATGCCTTGGAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAA 2014

Qy    598 ------------------------------------------AspLysAlaGluGluSer 603
Db   2015 AGAAGATCAACTCGCAGGAGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACA 2074

Qy    604 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 623
Db   2075 AAAGAGGACTTGAATGAGAATGTACCTCCAGTTTCCTTCTGGAGGATTCTGAAGCTGAAC 2134

Qy    624 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 643
Db   2135 TCAACTGAATGGCCTTATTTTGTGGTTGGTATATTTTGTGCTATTATAAACGGAGGCCTG 2194

Qy    644 HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAsp 662
Db   2195 CAACCAGCATTTTCAATAATATTTTCAAGGATTATAGGGATCTTTACCCGAGATGAGGAT 2254

Qy    663 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 682
```

```
Db   2255 CCTGAAACAAAACGACAGAATAGTAACATGTTTTCTGTATTGTTTCTAGTCCTTGGAATT 2314

Qy    683 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 702
Db   2315 ATTTCTTTTATTACATTTTTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTC 2374

Qy    703 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 722
Db   2375 ACTAAGCGGCTTCGATACATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTT 2434

Qy    723 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 742
Db   2435 GATGACCCTAAAAACACCACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAA 2494

Qy    743 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 762
Db   2495 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGG 2554

Qy    763 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 782
Db   2555 ACAGGCATTATTATATCCTTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATT 2614

Qy    783 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 802
Db   2615 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 2674

Qy    803 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 822
Db   2675 AAAGATAAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTC 2734

Qy    823 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 842
Db   2735 CGAACTGTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTG 2794

Qy    843 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 862
Db   2795 CAAGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATC 2854

Qy    863 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 882
Db   2855 ACCCAGGCAATGATGTATTTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTG 2914

Qy    883 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 902
Db   2915 GCAAATGAGTTCATGAACTTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGT 2974

Qy    903 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 922
Db   2975 GCCATGGCAGTGGGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCA 3034

Qy    923 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 942
Db   3035 GCAGCCCACGTCATCATGATCATTGAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCAC 3094

Qy    943 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 962
Db   3095 GGCCTCAAGCCAAATACGTTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTAT 3154

Qy    963 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 982
Db   3155 CCCACTCGACCAGACATCCCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAG 3214

Qy    983 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 1002
Db   3215 ACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAG 3274

Qy   1003 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 1022
Db   3275 CGCTTCTATGACCCCTTGGCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTG 3334

Qy   1023 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 1042
Db   3335 AATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGAC 3394

Qy   1043 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 1062
Db   3395 TGCAGCATTGCCGAGAACATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAG 3454

Qy   1063 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 1082
Db   3455 ATTATGCAGGCAGCCAAGGAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAA 3514

Qy   1083 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1102
Db   3515 TACAACACCAGAGTAGGAGACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATT 3574

Qy   1103 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 1122
Db   3575 GCCATAGCTCGCGCTCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCTACATCA 3634

Qy   1123 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1142
Db   3635 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 3694

Qy   1143 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1162
Db   3695 CGCACCTGCATTGTGATCGCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTG 3754

Qy   1163 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1182
Db   3755 GTGTTTCAGAATGGCAAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGGCCCAGAAA 3814

Qy   1183 AspIleTyrPheLysLeuValAsnAlaGlnSer 1193
Db   3815 GGCATCTATTTTTCCATGGTCAGTGTCCAGGCT 3847
```

RESULT 9
AAD03489

```
ID   AAD03489 standard; cDNA; 4317 BP.
XX
AC   AAD03489;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) cDNA #2.
XX
KW   Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             70..3912
FT                   /*tag=  a
FT                   /product= "Dog P-glycoprotein (PGP) #2"
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00304.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
```

```
PS   Claim 1; Page 66-72; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) cDNA. The
CC   PGP enzyme functions as an efflux pump exporting small molecules
CC   across the cell membrane. This enzyme is a member of the ABC
CC   transporter family.
XX
SQ   Sequence 4317 BP; 1293 A; 844 C; 1019 G; 1161 T; 0 other;

Alignment Scores:
Pred. No.:              0          Length:         4317
Score:                  3549.00    Matches:        680
Percent Similarity:     74.65%     Conservative:   239
Best Local Similarity:  55.24%     Mismatches:     270
Query Match:            59.03%     Indels:         42
DB:                     22         Gaps:           8

US-09-873-409-6 (1-1195) x AAD03489 (1-4317)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    217 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 276

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
Db    277 GTTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 336

Qy     35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
Db    337 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 396

Qy     45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db    397 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 456

Qy     60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db    457 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 516

Qy     80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db    517 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 576

Qy    100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db    577 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 636

Qy    119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db    637 AAAGTTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 696

Qy    139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db    697 ACACCTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 756

Qy    159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db    757 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 816

Qy    179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db    817 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 876

Qy    199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
Db    877 GGACAAAAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGG 936

Qy    219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238
Db    937 ATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCA 996

Qy    239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
Db    997 TCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TAT 1050

Qy    259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
Db   1051 ACTATTGGACAGGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGA 1110

Qy    279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
Db   1111 CAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTC 1170

Qy    299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
Db   1171 AAGATAATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGAT 1230

Qy    319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
Db   1231 AATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAA 1290

Qy    339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
Db   1291 GTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTT 1350

Qy    359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
Db   1351 GGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCC 1410

Qy    379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
Db   1411 ACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTT 1470

Qy    399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
Db   1471 CGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAA 1530

Qy    419 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
Db   1531 AACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAA 1590

Qy    439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
Db   1591 GCCAATGCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAG 1650

Qy    459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
Db   1651 AGAGGGGCCCGGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTT 1710

Qy    479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
Db   1711 CGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAA 1770

Qy    499 SerAlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAla 518
Db   1771 GCAGTGGTTCAGGTGGCCCTGGATAAGGCCAGAAAAGGCCGGACTACCATTGTGATAGCT 1830

Qy    519 HisArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeu 538
Db   1831 CATCGTTTGTCTACAGTTCGTAATGCCGATGTCATTGCTGGTTTTGATGATGGAGTCATT 1890

Qy    539 AlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuVal 558
```

```
Db   1891 GTGGAGAAAGGAAATCATGATGAACTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTC 1950

Qy    559 MetSerGlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGlu 578
Db   1951 ACAATGCAG---ACAAGAGGAAATGAAATTGAGTTAGAAAATGCCACTGGTGAATCCAAA 2007

Qy    579 ArgLysThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIle--- 597
Db   2008 AGTGAAAGTGATGCCTTGGAAATG---TCTCCAAAAGATTCAGGGTCCAGTTTAATAAAA 2064

Qy    598 ----------------------------------------AspLysAlaGluGluSer 603
Db   2065 AGAAGATCAACTCGCAGGAGTATACATGCACCACAAGGCCAAGACAGAAAGCTTGGTACA 2124

Qy    604 ThrGlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsn 623
Db   2125 AAAGAGGACTTGAATGAGAATGTTCCTTCAGTTTCCTTCTGGAGGATTCTGAAGCTGAAC 2184

Qy    624 LysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrVal 643
Db   2185 TCAACTGAATGGCCTTATTTTGTGGTTGGTATATTTTGTGCTATTATAAACGGAGGCCTG 2244

Qy    644 HisProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAsp 662
Db   2245 CAACCAGCATTTTCAATAATATTTTCAAGGATTATAGGGATCTTTACCCGAGATGAGGAT 2304

Qy    663 LysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyVal 682
Db   2305 CCTGAAACAAAACGACAGAATAGTAACATGTTTTCTGTATTGTTTCTAGTCCTTGGAATT 2364

Qy    683 IleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeu 702
Db   2365 ATTTCTTTTATTACATTTTTCCTCCAGGGCTTCACATTTGGCAAAGCTGGGGAGATCCTC 2424

Qy    703 ThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPhe 722
Db   2425 ACTAAGCGGCTTCGATACATGGTTTTCAGATCCATGCTGAGACAGGATGTCAGCTGGTTT 2484

Qy    723 AspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGln 742
Db   2485 GATGACCCTAAAAACACCACTGGAGCATTGACAACCAGGCTTGCCAATGATGCGGCTCAA 2544

Qy    743 IleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGly 762
Db   2545 GTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTCATTACCCAGAATATAGCAAATCTTGGG 2604

Qy    763 LeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIle 782
Db   2605 ACAGGCATTATTATATCCTTAATCTATGGTTGGCAATTAACACTTTTACTCTTAGCAATT 2664

Qy    783 AlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsn 802
Db   2665 GTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTG 2724

Qy    803 LysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIle 822
Db   2725 AAAGATAAGAAAGAGCTAGAAGGAGCTGGGAAGATTGCTACAGAAGCCATCGAAAACTTC 2784

Qy    823 ArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeu 842
Db   2785 CGAACTGTTGTTTCTTTGACTCGGGAGCAGAAGTTTGAATACATGTATGCACAGAGTTTG 2844

Qy    843 GlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPhe 862
Db   2845 CAAGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTCGGGGTCTCATTTTCTATC 2904

Qy    863 SerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIle 882
Db   2905 ACCCAGGCAATGATGTATTTTTCCTATGCTGGCTGTTTCCGGTTTGGTGCCTACTTGGTG 2964

Qy    883 GlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGly 902
Db   2965 GCAAATGAGTTCATGAACTTTCAGGATGTTCTTTTGGTATTCTCAGCTATTGTCTTTGGT 3024

Qy    903 AlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGly 922
Db   3025 GCCATGGCAGTGGGGCAGGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCA 3084

Qy    923 AlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGlu 942
Db   3085 GCAGCCCACGTCATCATGATCATTGAAAAAAGCCCTCTGATTGACAGCTACAGCCCTCAC 3144

Qy    943 GlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyr 962
Db   3145 GGCCTCAAGCCAAATACGTTGGAAGGAAATGTGACATTTAATGAGGTCGTGTTCAACTAT 3204

Qy    963 ProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLys 982
Db   3205 CCCACTCGACCAGACATCCCCGTGCTCCAGGGGCTGAGCCTCGAGGTGAAGAAGGGCCAG 3264

Qy    983 ThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGln 1002
Db   3265 ACGCTGGCCCTCGTAGGTAGCAGTGGCTGTGGGAAGAGCACAGTTGTTCAGCTCCTAGAG 3324

Qy   1003 ArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeu 1022
Db   3325 CGCTTCTATGACCCCTTGGCTGGTTCAGTGCTAATTGATGGCAAAGAGATAAAGCACCTG 3384

Qy   1023 AsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsn 1042
Db   3385 AATGTCCAGTGGCTCCGAGCACACCTGGGCATCGTGTCTCAGGAGCCCATCCTGTTTGAC 3444

Qy   1043 CysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGlu 1062
Db   3445 TGCAGCATTGCCGAGAACATTGCCTATGGAGACAACAGCCGGGTCGTATCACATGAAGAG 3504

Qy   1063 IleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLys 1082
Db   3505 ATTATGCAGGCAGCCAAGGAGGCCAACATACACCACTTCATCGAGACACTCCCTGAGAAA 3564

Qy   1083 TyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeu 1102
Db   3565 TACAACACCAGAGTAGGAGACAAAGGAACCCAGCTCTCTGGTGGCCAGAAACAGCGCATT 3624

Qy   1103 AlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSer 1122
Db   3625 GCCATAGCTCGCGCTCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCTACATCA 3684

Qy   1123 AlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGly 1142
Db   3685 GCTCTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGC 3744

Qy   1143 ArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleVal 1162
Db   3745 CGCACCTGCATTGTGATCGCCCACCGCTTGTCCACCATCCAGAATGCAGATTTAATAGTG 3804

Qy   1163 ValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArg 1182
Db   3805 GTGTTTCAGAATGGCAAAGTCAAGGAGCATGGCACACATCAACAGCTGCTGGCTCAGAAA 3864

Qy   1183 AspIleTyrPheLysLeuValAsnAlaGlnSer 1193
Db   3865 GGCATCTATTTTTCCATGATCAGTGTCCAGGCT 3897
```

```
RESULT 10
AAF86127
ID   AAF86127 standard; cDNA; 4186 BP.
XX
AC   AAF86127;
XX
DT   25-JUN-2001  (first entry)
XX
DE   Cynomologous monkey P-glycoprotein cDNA.
XX
KW   Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   efflux pump; ss.
```

```
XX
OS   Macaca fascicularis.
XX
FH   Key             Location/Qualifiers
FT   CDS             100..3942
FT                   /*tag=  a
FT                   /product= "PGP"
FT                   /note= "P-glycoprotein"
XX
PN   WO200123565-A1.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26592.
XX
PR   28-SEP-1999;   99US-0156921.
PR   12-OCT-1999;   99US-0158818.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-Crespi DT,  Crespi CL;
XX
DR   WPI; 2001-316136/33.
DR   P-PSDB; AAB81064.
XX
PT   Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein
PT   (PGP) and homologous PGP polypeptides are useful for predicting
PT   bioavailability of compound and increasing PGP transporter activity in
PT   cell  -
XX
PS   Example 1; Page 51-57; 84pp; English.
XX
CC   This invention relates to a polynucleotide sequence encoding a
CC   cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the
CC   PGP protein. PGP, also known as multidrug transporter, MDR1 is a member
CC   of the ABC transporter superfamily. The enzyme serves as an efflux pump
CC   exporting small molecules across the cell membrane. The invention
CC   includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence
CC   and protein, and also that of an allelic variant. The PGP polynucleotide
CC   sequence is useful for increasing PGP transporter activity in a cell.
CC   Antisense sequences of the cDNA are useful for inhibiting PGP transport
CC   activity in a mammalian cell. They may also be used for increasing the
CC   bioavailability of a drug. The present sequence represents cDNA encoding
CC   cynomologous monkey P-glycoprotein.
XX
SQ   Sequence 4186 BP; 1226 A; 799 C; 1039 G; 1122 T; 0 other;

Alignment Scores:
Pred. No.:              0           Length:         4186
Score:                  3548.00     Matches:        679
Percent Similarity:     75.37%      Conservative:   248
Best Local Similarity:  55.20%      Mismatches:     261
Query Match:            59.02%      Indels:         42
DB:                     22          Gaps:           10

US-09-873-409-6 (1-1195) x AAF86127 (1-4186)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
             |||::::::|||   |||||::::::::::::|||||   |||||||||||    |||
Db    250 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGAGCTGGACTTCCTCTCATGATGCTG 309

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSer---------------------GlyCys 33
             |||   |||:::|||:::|||            :::              |||
Db    310 GTGTTTGGAGACATGACGGATACCTTTGCAAATGCAGGAAATTTAGGAGATTTAGGAGCT 369

Qy     34 LeuValGlnThrAsnThrTyrSer----------------------------------- 41
             |||:::   ||||||:::    :::
Db    370 CTGTTG---ACTAATAGCAGTAATATCACTGATACAGTGCCCGTCATGAATCTGGAGGAA 426

Qy     42 ---PhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
                      |||          ||||||   |||||||||          |||:::
Db    427 GATATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCT 486

Qy     61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
             ||||||||:::|||   |||   :::   |||   ||||||      :::||||||||
Db    487 TACATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAA 546

Qy     81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
             ||||||||||||:::::::::   |||:::|||||||||||||||      |||:::|||
Db    547 CAGTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGG 606

Qy    101 GluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLys 119
             |||||||||||||||:::|||   |||:::   ||||||::::::||||||||||||||
Db    607 GAGCTTAACACCCGGCTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAA 666

Qy    120 IleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuVal 139
             |||   :::   ||||||:::|||:::||||||         |||      ||||||
Db    667 ATTGGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACA 726

Qy    140 LysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAla 159
             :::||||||||||||||||||||||   |||:::   ||||||::::::      ||||||
Db    727 CGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCT 786

Qy    160 AlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLys 179
             |||      ::::::::::::   |||   |||      |||||||   ||||||:::|||
Db    787 GCAGTCTGGGCAAAGATACTGTCTTCATTTACTGATAAAGAACTCTTAGCTTATGCAAAA 846

Qy    180 AlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAla 199
             ||||||||||||||||||||||||||||::::::|||||||||||||||||||||
Db    847 GCTGGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGA 906

Qy    200 GlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIle 219
             |||:::|||||||||:::||||||   :::||||||::::::||||||      ||||||
Db    907 CAAAAGAAAGAACTCGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATA 966

Qy    220 LysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThr 239
             |||:::   |||   :::   :::|||:::||||||   :::      :::      :::
Db    967 AAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTTATCTATGCATCT 1026

Qy    240 TyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThr 259
             |||   |||||||||||||||||||||:::|||:::|||:::   |||      |||:::
Db   1027 TATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAAAGGAA------TATTCT 1080

Qy    260 IleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAla 279
             ||||||   ||||||   |||||||||||||||:::      ::::::   :::|||
Db   1081 ATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAG 1140

Qy    280 AlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGln 299
             |||   |||      |||   ||||||   |||||||||||||||||||   ||||||:::
Db   1141 GCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTTTGAAATCTTCAAG 1200

Qy    300 ValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSer 319
             :::||||||   |||||||||||||||::::::|||   :::|||:::||||||::::::
Db   1201 ATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAAT 1260

Qy    320 IleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIle 339
             |||:::|||   :::||||||:::||||||   |||:::||||||||||||   :::
Db   1261 ATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTT 1320

Qy    340 LysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGly 359
             |||||||||||||||||||||||||::::::::::|||||||:::||||||||||||||||
Db   1321 AAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGA 1380

Qy    360 LeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAsp 379
             :::|||   ||||||||||||   |||||||||:::||||||||||||||||||
Db   1381 AACAGCGGCTGTGGGAAGAGCACAACGGTCCAGCTGATGCAGAGGCTTTATGACCCCACA 1440

Qy    380 AspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArg 399
             :::|||   :::   |||||      ||||||||||||   :::|||||||||      |||
Db   1441 GAGGGCATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAACGTAAGGTTTCTACGG 1500

Qy    400 AspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsn 419
             :::   ||||||||||||||||||||||||||||||   |||||||||:::   |||
Db   1501 GAAATCATCGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAAC 1560

Qy    420 IleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAla 439
```

```
Db   1561 ATTCGCTATGGTCGTGAAGATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCC 1620

Qy    440 AsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLys 459
Db   1621 AATGCCTATGACTTTATCATGAAACTGCCTCAGAAATTTGACACCCTGGTTGGAGAGAGA 1680

Qy    460 GlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArg 479
Db   1681 GGGGCCCAGCTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGC 1740

Qy    480 AsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSer 499
Db   1741 AACCCCAAGATCCTCCTGCTGGACGAGGCCACGTCAGCCTTGGACACAGAAAGTGAAGCA 1800

Qy    500 AlaValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHis 519
Db   1801 GTGGTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCAT 1860

Qy    520 ArgLeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAla 539
Db   1861 CGTTTGTCTACGGTTCGTAATGCCGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTG 1920

Qy    540 GluLysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMet 559
Db   1921 GAGAAAGGAAATCATGATGAGCTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACA 1980

Qy    560 SerGlnAspIle---------------------LysLysAlaAspGlu---------Gln 569
Db   1981 ATGCAGACAGCAGGAAATGAAATTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAA 2040

Qy    570 MetGluSerMetThrTyrSerThrGluArgLysThrAsnSerLeu--------------- 584
Db   2041 ATTGATACCTTGGAAATGTCTTCACATGATTCAGGATCCAGTCTAATAAGAAAAAGATCC 2100

Qy    585 ProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr 604
Db   2101 ACTCGTAGGAGTGTCCGTGGATCACAAGGCCAA------GACAGAAAGCTTAGTACCAAA 2154

Qy    605 GlnSerLysGluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLys 624
Db   2155 GAGGCTCTGGATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTA 2214

Qy    625 ProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHis 644
Db   2215 ACTGAGTGGCCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGTCTGCAA 2274

Qy    645 ProValPheSerIleIlePheAlaLysIleIleThrMetPhe---GlyAsnAsnAspLys 663
Db   2275 CCAGCATTTGCAGTAATATTTTCAAAGATTATAGGGATTTTTACAAGAAATGATGATGCC 2334

Qy    664 ThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIle 683
Db   2335 GAAACAAAACGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGTCCTTGGAATTGTT 2394

Qy    684 CysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThr 703
Db   2395 TCTTTTATTACATTTTTCCTTCAGGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACC 2454

Qy    704 MetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAsp 723
Db   2455 AAGCGGCTCCGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGCTGGTTTGAT 2514

Qy    724 GluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIle 743
Db   2515 GACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTT 2574

Qy    744 GlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeu 763
Db   2575 AAAGGGGCTATAGGTTCCAGGCTTGCTATAATTACCCAGAATATAGCAAATCTTGGGACA 2634

Qy    764 SerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAla 783
Db   2635 GGAATAATTATATCCTTAATCTATGGTTGGCAACTGACACTGTTACTCTTAGCAATTGTA 2694

Qy    784 ProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLys 803
Db   2695 CCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAA 2754

Qy    804 AspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArg 823
Db   2755 GATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGA 2814

Qy    824 ThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGln 843
Db   2815 ACTGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGATCAGAGTTTGCAG 2874

Qy    844 ThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSer 863
Db   2875 GTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGAATCACGTTTTCCTTCACG 2934

Qy    864 HisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGln 883
Db   2935 CAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCA 2994

Qy    884 AlaGlyArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAla 903
Db   2995 CATAGTCTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCC 3054

Qy    904 MetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAla 923
Db   3055 ATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAGTATCAGCA 3114

Qy    924 AlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGly 943
Db   3115 GCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACAGAAGGC 3174

Qy    944 LysLysProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrPro 963
Db   3175 CTAAAGCCGAACACATTGGAAGGAAATGTCACATTTAATGAAGTTGTATTCAACTATCCC 3234

Qy    964 CysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThr 983
Db   3235 ACCCGACTGGACATCCCAGTGCTTCAGGGGCTGAGCCTGGAAGTGAAGAAGGGCCAGACG 3294

Qy    984 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 1003
Db   3295 CTGGCCCTGGTGGGCAGCAGTGGCTGTGGGAAGAGCACGGTGGTCCAGCTCCTGGAGCGG 3354

Qy   1004 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 1023
Db   3355 TTCTATGACCCCTTGGCGGGGAAAGTGCTGCTTGACGGCAAAGAAATAAAGCAACTGAAT 3414

Qy   1024 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 1043
Db   3415 GTTCAGTGGCTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGC 3474

Qy   1044 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 1063
Db   3475 AGCATTAGTGAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATC 3534

Qy   1064 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 1083
Db   3535 GTGAGGGCAGCCAAGGAGGCCAATATACACGCCTTCATCGAGTCACTGCCTAATAAATAT 3594

Qy   1084 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 1103
Db   3595 AGCACCAGAGTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCC 3654

Qy   1104 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 1123
Db   3655 ATAGCTCGTGCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACATCAGCT 3714

Qy   1124 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 1143
Db   3715 CTGGATACAGAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGT 3774
```

```
Qy    1144 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValVal 1163
           |||||:::|||:::    ||||||||||   |||||||||||||||||||||||||
Db    3775 ACCTGCATTGTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTG 3834

Qy    1164 LeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAsp 1183
                   |||||::::::|||||   ||||||||||:::||||||       :::
Db    3835 TTTCAGAATGGCAGAGTCAAGGAGCACGGCACACATCAGCAGCTGCTGGCACAGAAAGGC 3894

Qy    1184 IleTyrPheLysLeuValAsnAlaGlnSer 1193
           ||||||||   :::|||:::   |||:::
Db    3895 ATCTATTTTTCAATGGTCAGTGTCCAGGCT 3924

RESULT 11
AAZ49333
ID   AAZ49333 standard; cDNA; 3860 BP.
XX
AC   AAZ49333;
XX
DT   14-MAR-2000  (first entry)
XX
DE   Human G185V mutant multidrug resistance-1 (MDR-1) cDNA.
XX
KW   Multidrug resistance; MDR-1; P-glycoprotein;
KW   transmembrane efflux pump; haematopoietic stem cell; transduction;
KW   bone marrow transplantation; chemotherapy; radiation therapy; cancer;
KW   gene therapy; gene replacement; genetic defect; thalassaemia;
KW   Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
KW   cytokine; mutant; ds.
XX
OS   Synthetic.
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3843
FT                   /*tag=  a
FT                   /product= "Human G185V mutant MDR-1 protein"
XX
PN   WO9961589-A2.
XX
PD   02-DEC-1999.
XX
PF   27-MAY-1999;   99WO-US11825.
XX
PR   28-MAY-1998;   98US-0086988.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI   Sorrentino B,  Bunting K;
XX
DR   WPI; 2000-072615/06.
DR   P-PSDB; AAY58187.
XX
PT   Ex vivo expansion of hematopoietic stem cells transduced with a
PT   sequence encoding human multidrug resistance-1, used for bone marrow
PT   transplantation -
XX
PS   Example 1; Page 79-82; 113pp; English.
XX
CC   This sequence represents cDNA encoding human G185V mutant multidrug
CC   resistance protein MDR-1, where the the Gly residue at postion 185
CC   of the wild-type protein (AAY58186) is replaced by Val. MDR-1 is a
CC   transmembrane efflux pump, responsible for the export of drugs from
CC   cells, particularly cancer cells. The wild-type MDR-1 shows increased
CC   resistance to etoposide and decreased resistance to vinca alkaloids
CC   compared with the G185V mutant. The invention relates to transducing
CC   haematopoietic stem cells with nucleic acid encoding an MDR protein
CC   and culturing the modified cells. The modified haematopoietic stem
CC   cells are useful in bone marrow transplantation (to reconstitute
CC   haematopoietic systems in patients who have undergone chemotherapy or
CC   radiation therapy) and in ex vivo gene therapy of genetic defects in
CC   cells derived from haematopoietic stem cells, e.g., thalassaemia,
CC   Gauchier's disease, sickle cell anaemia or leukaemia. The modified
CC   cells can also be used to identify factors involved in regulating
CC   proliferation and differentiation in haematopoietic stem cells.
CC   Haematopoietic stem cells that express MDR-1 will be protected against
CC   chemotherapeutic agents, so can be engrafted while the patient is
CC   undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells
CC   provides sufficient cells to permit standard biochemical analysis.
CC   Overexpression of MDR-1 allows cytokine-driven expansion of
CC   haematopoietic stem cells by at least 10-fold compared with a maximum
CC   of 4-fold in known procedures.
XX
SQ   Sequence 3860 BP; 1135 A; 746 C; 956 G; 1023 T; 0 other;

Alignment Scores:
Pred. No.:              0              Length:         3860
Score:                  3545.00        Matches:        681
Percent Similarity:     75.55%         Conservative:   246
Best Local Similarity:  55.50%         Mismatches:     264
Query Match:            58.97%         Indels:         36
DB:                     21             Gaps:           9

US-09-873-409-6 (1-1195) x AAZ49333 (1-3860)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||::::::|||   ||||||::::::::::::||||||   |||||||||||   |||
Db     151 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210

Qy      21 ValLeuGlyGluMetSerAsp-------------------------AsnLeuIleSer 31
           |||   |||||||||:::|||                          :::|||:::|||
Db     211 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 270

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe----------------- 43
                       ::::::      |||        ||||||
Db     271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                 |||         ||||||   |||||||||         |||:::      |||
Db     331 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 390

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
           ||||||:::|||   |||   :::   |||   ||||||      :::||||||||||||
Db     391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
           |||||||||::::::::      |||:::|||||||||||||||       |||:::||||||
Db     451 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 510

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
           ||||||||||||:::|||   |||:::   ||||||::::::   |||||||||||||||
Db     511 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 570

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPhePheSerIleGlyLeuAlaValGlyLeuValLys 140
              :::   ||||||:::|||:::||||||         |||         ||||||      :::
Db     571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 630

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
           ||||||||||||||||||   |||:::   ||||||::::::      |||||||||
Db     631 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
              ::::::::::::   |||   |||   |||||||||   ||||||:::|||||||
Db     691 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
           ||||||||||||||||||||||||::::::||||||||||||||||||||||      |||
Db     751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810

Qy     201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
           :::|||||||||:::||||||   :::||||||::::::||||||      |||||||||
Db     811 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 870

Qy    -221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
           :::   |||   :::   :::|||:::||||||   :::         :::      :::|||
Db     871 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 930
```

```
Qy   241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db   931 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 984

Qy   261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db   985 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1044

Qy   281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db  1045 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104

Qy   301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db  1105 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1164

Qy   321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db  1165 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1224

Qy   341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db  1225 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1284

Qy   361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db  1285 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1344

Qy   381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db  1345 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1404

Qy   401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db  1405 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1464

Qy   421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db  1465 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1524

Qy   441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db  1525 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1584

Qy   461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db  1585 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1644

Qy   481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db  1645 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1704

Qy   501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db  1705 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 1764

Qy   521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db  1765 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 1824

Qy   541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db  1825 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 1884

Qy   561 Gln-----------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
Db  1885 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 1944

Qy   575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 589
Db  1945 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2004

Qy   590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
Db  2005 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2064

Qy   608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627
Db  2065 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2124

Qy   628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
Db  2125 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2184

Qy   648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
Db  2185 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2244

Qy   667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
Db  2245 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2304

Qy   687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
Db  2305 ACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2364

Qy   707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
Db  2365 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2424

Qy   727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
Db  2425 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2484

Qy   747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
Db  2485 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2544

Qy   767 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
Db  2545 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 2604

Qy   787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
Db  2605 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 2664

Qy   807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
Db  2665 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 2724

Qy   827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
Db  2725 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2784

Qy   847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
Db  2785 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 2844

Qy   867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
Db  2845 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 2904

Qy   887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
Db  2905 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 2964

Qy   907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
Db  2965 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3024

Qy   927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
Db  3025 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3084

Qy   947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
```

```
              :::|||  |||||||||:::  |||  ||||||  |||  ||||||  ||||||
Db     3085 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3144

Qy      967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
              |||:::   :::||||:::||||||||||   :::::::::|||:::|||:::|||
Db     3145 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3204

Qy      987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
              ||||||||||||||||||||||||||||||||||||   ||||||||||:::|||  ||||||
Db     3205 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3264

Qy     1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
              |||:::   |||:::||||||  ||||||  :::  |||  ||||||||||||||||
Db     3265 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3324

Qy     1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
              ||||||:::  :::  ||||||  |||||||||:::||||||:::||||||||||||
Db     3325 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3384

Qy     1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
              ||||||||||||||||||||||||||||||||||||      :::||||||      |||
Db     3385 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3444

Qy     1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
              |||      ||||||||||||:::||||||||||  ||||||  ||||||:::|||:::
Db     3445 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3504

Qy     1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
              ||||||  ||||||  ||||||||||||||||||||||||||||||:::||||||||||||
Db     3505 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3564

Qy     1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
              ||||||:::::::::|||  ||||||||||||||||||||||||||||||||||||
Db     3565 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 3624

Qy     1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
              :::||||||||||||||||||  ||||||||||||||||||  ||||||||||||:::
Db     3625 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 3684

Qy     1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
              |||:::  ||||||||||||  ||||||||||||||||||||||||||||      |||
Db     3685 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 3744

Qy     1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
              |||::::::||||||  ||||||||||||:::||||||      :::  |||||||||
Db     3745 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 3804

Qy     1187 LysLeuValAsnAlaGlnSer 1193
             .:::|||:::  |||:::
Db     3805 TCAATGGTCAGTGTCCAGGCT 3825
```

```
RESULT 12
ABA94366
ID   ABA94366 standard; DNA; 3860 BP.
XX
AC   ABA94366;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human BCRP DNA related seq Id No. 3.
XX
KW   Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;
KW   hepatotropic; neurodegenerative; cytostatic; antianemic; muscular; BCRP;
KW   cardiant; gene therapy; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3843
FT                   /*tag=  a
XX
PN   WO200192877-A2.
XX
PD   06-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US17459.
XX
PR   31-MAY-2000; 2000US-0584586.
PR   29-MAY-2001; 2001US-0866866.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI   Sorrentino B,  Schuetz J;
XX
DR   WPI; 2002-114368/15.
DR   P-PSDB; ABB07267.
XX
PT   Identifying a stem cell, for treating e.g., muscular dystrophy,
PT   myocardial infarction, Parkinson's disease, or neurodegenerative
PT   disorders, comprises detecting the expression of an ATP transport
PT   protein (BCRP) by a cell  -
XX
PS   Disclosure; Page 59-60; 87pp; English.
XX
CC   The invention provides a method of identifying and/or isolating a stem
CC   cell that involves detecting the expression of an ATP transport protein
CC   containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
CC   comprising stem cells. The isolated stem cells may be used in the
CC   treatment of diseases such as muscular dystrophy, degenerative liver
CC   disorder, myocardial infarction, Parkinson's disease, degenerative
CC   disorders of the brain, and for tissue regeneration or replacement.
CC   Haematopoietic cells can be used in bone marrow transplants (e.g., for
CC   treatment of leukemia) and for ex vivo gene therapy for treating blood
CC   diseases such as sickle cell anemia and thalassemia. The stem cells can
CC   also be used as cell targets in gene therapy protocols. The present
CC   sequence represents a sequence related to the BCRP for which no relevant
CC   information has been provided in the specification.
XX
SQ   Sequence 3860 BP; 1135 A; 746 C; 956 G; 1023 T; 0 other;

Alignment Scores:
Pred. No.:              0              Length:         3860
Score:                  3545.00        Matches:        681
Percent Similarity:     75.55%         Conservative:   246
Best Local Similarity:  55.50%         Mismatches:     264
Query Match:            58.97%         Indels:         36
DB:                     24             Gaps:           9

US-09-873-409-6 (1-1195) x ABA94366 (1-3860)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
              |||::::::|||  ||||||:::::::::::::||||||  ||||||||||||  |||
Db      151 ATGGTGGTGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210

Qy       21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
              |||  |||||||||:::|||                           :::|||:::|||
Db      211 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 270

Qy       32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                              ::::::  |||      ||||||
Db      271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330

Qy       44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                  |||       ||||||  |||||||||           |||:::       |||
Db      331 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 390

Qy       62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
              ||||||:::|||  |||  :::  |||  ||||||       :::||||||||||||
Db      391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450

Qy       82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
              ||||||||:::::::::  |||:::|||||||||||||||      |||:::||||||
Db      451 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 510

Qy      102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
              ||||||||||||:::|||  |||:::  ||||||::::::  ||||||||||||||
```

```
Db    511 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 570

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 630

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    631 GGTTGGAAGCTAACCCTTGTGATTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    691 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db    811 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 870

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db    871 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 930

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db    931 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 984

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db    985 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1044

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db   1045 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db   1105 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1164

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db   1165 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1224

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   1225 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1284

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   1285 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1344

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   1345 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1404

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   1405 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1464

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1465 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1524

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1525 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1584

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   1585 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1644

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   1645 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1704

Qy    501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db   1705 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 1764

Qy    521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db   1765 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 1824

Qy    541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db   1825 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 1884

Qy    561 Gln-----------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
Db   1885 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 1944

Qy    575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 589
Db   1945 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2004

Qy    590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
Db   2005 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2064

Qy    608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627
Db   2065 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2124

Qy    628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
Db   2125 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2184

Qy    648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
Db   2185 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2244

Qy    667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
Db   2245 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2304

Qy    687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
Db   2305 ACATTTTTCCTTCAAGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2364

Qy    707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
Db   2365 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2424

Qy    727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
Db   2425 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2484

Qy    747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
Db   2485 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2544

Qy    767 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
Db   2545 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 2604

Qy    787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
Db   2605 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 2664

Qy    807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
Db   2665 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 2724
```

```
Qy    827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
          ||||||||:::|||:::  ||||||  ||||||   :::   ||||||        :::
Db   2725 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2784

Qy    847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
          ||||||:::   :::||||||  |||   |||     ::::::|||:::   |||
Db   2785 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 2844

Qy    867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
          :::||||||:::||||||      ||||||||||||||||||||||:::
Db   2845 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 2904

Qy    887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
          |||:::   |||    :::   :::||||||:::|||:::   :::||||||||||||:::
Db   2905 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 2964

Qy    907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
          |||:::            ||||||:::|||:::||||||||      |||||||||:::
Db   2965 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3024

Qy    927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
                ::::::||||||   |||   |||||||||   ||||||          |||
Db   3025 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3084

Qy    947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
          :::|||   |||||||||:::   |||   ||||||   |||   ||||||   ||||||
Db   3085 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3144

Qy    967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
          |||:::   :::|||:::||||||||||||   :::::::::|||:::|||:::|||
Db   3145 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3204

Qy    987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
          ||||||||||||||||||||||||||||||   ||||||||||||:::|||   ||||||
Db   3205 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3264

Qy   1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
          |||:::   |||:::||||||   ||||||      :::   |||   ||||||||||||||
Db   3265 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3324

Qy   1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
          ||||||:::   :::   ||||||   |||||||||:::||||||:::||||||||||||
Db   3325 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3384

Qy   1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
          ||||||||||||||||||||||||||||||||||||         :::||||||     |||
Db   3385 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3444

Qy   1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
          |||      ||||||||||||:::|||||||||   ||||||   ||||||:::|||:::
Db   3445 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3504

Qy   1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
          ||||||   ||||||   ||||||||||||||||||||||||||||:::|||||||||||||
Db   3505 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3564

Qy   1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
          ||||||:::::::::|||   ||||||||||||||||||||||||||||||||||||
Db   3565 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 3624

Qy   1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
          :::||||||||||||||||||   ||||||||||||||||||   ||||||||||||:::
Db   3625 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 3684

Qy   1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
          |||:::   ||||||||||||   |||||||||||||||||||||||||||       |||
Db   3685 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 3744

Qy   1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
          |||::::::||||||   ||||||||||||:::||||||       :::   |||||||||
Db   3745 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 3804

Qy   1187 LysLeuValAsnAlaGlnSer 1193
          :::|||:::   |||:::
Db   3805 TCAATGGTCAGTGTCCAGGCT 3825
```

RESULT 13
AAH57442
ID   AAH57442 standard; cDNA; 4349 BP.
XX
AC   AAH57442;
XX
DT   10-SEP-2001 (first entry)
XX
DE   Human intestine cell specific cDNA sequence SEQ ID NO:282.
XX
KW   Human; tissue specific; diagnosis; brain; heart; skeletal muscle;
KW   lung; liver; uterus; ovary; stomach; intestine; kidney; pancreas; ss;
KW   metabolic disease; developmental disease; cytostatic; immunomodulatory;
KW   neuroprotective; gene therapy; cancer; immunopathology; neuropathology.
XX
OS   Homo sapiens.
XX
PN   WO200132927-A2.
XX
PD   10-MAY-2001.
XX
PF   02-NOV-2000; 2000WO-US30396.
XX
PR   04-NOV-1999;   99US-0163508.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Sornasse T, Seilhamer JJ, Watson GA;
XX
DR   WPI; 2001-291057/30.
XX
PT   New cell and tissue specific polynucleotides useful for diagnosis,
PT   prognosis or monitoring of treatments for disorders where the gene is
PT   associated with a cancer, immunopathology or neuropathology -
XX
PS   Claim 1; Page 207-208; 327pp; English.
XX
CC   AAH57161 to AAH57576 represent cell and tissue specific polynucleotide
CC   sequences (I). (I) can have cytostatic, immunomodulatory and
CC   neuroprotective activities, and can be used in gene therapy. (I) and
CC   proteins (II) encoded by then are used in high throughput screening
CC   assays to select DNA molecules, RNA molecules, peptide nucleic acids,
CC   mimetics, peptides, proteins, agonists, antagonists, antibodies or
CC   their fragments, immunoglobulins, inhibitors, drug compounds and
CC   pharmaceutical agents. Expression of (I) in a sample indicates the
CC   differentiation of embryonic stem cells into a tissue selected from
CC   brain, heart, kidney, liver, lung, skeletal muscle or pancreatic
CC   tissues. (I) and (II) are used to produce an expression profile that
CC   defines a metabolic or developmental process, treatment, condition,
CC   disease or disorder. The gene profile can be used for diagnosis,
CC   prognosis or monitoring of treatments and for investigating a
CC   predisposition to a disorder where the gene is associated with a
CC   cancer, immunopathology or neuropathology.
XX
SQ   Sequence 4349 BP; 1304 A; 816 C; 1048 G; 1181 T; 0 other;

Alignment Scores:

Pred. No.:	0	Length:	4349
Score:	3545.00	Matches:	681
Percent Similarity:	75.55%	Conservative:	246
Best Local Similarity:	55.50%	Mismatches:	264
Query Match:	58.97%	Indels:	36
DB:	22	Gaps:	9

US-09-873-409-6 (1-1195) x AAH57442 (1-4349)

```
Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          |||:::::::|||   ||||||:::::::::::::||||||   ||||||||||||   |||
Db    276 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 335
```

```
Qy     21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db    336 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 395

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe----------------- 43
Db    396 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 455

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    456 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 515

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    516 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 575

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    576 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 635

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    636 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 695

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    696 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 755

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    756 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 815

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    816 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 875

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    876 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 935

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db    936 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 995

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db    996 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1055

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db   1056 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1109

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db   1110 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1169

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db   1170 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1229

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db   1230 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1289

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db   1290 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1349

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   1350 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1409

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   1410 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1469

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   1470 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1529

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   1530 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1589

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1590 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1649

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1650 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1709

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   1710 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1769

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   1770 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1829

Qy    501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db   1830 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 1889

Qy    521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db   1890 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 1949

Qy    541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db   1950 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 2009

Qy    561 Gln-----------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
Db   2010 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 2069

Qy    575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 589
Db   2070 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2129

Qy    590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
Db   2130 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2189

Qy    608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627
Db   2190 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2249

Qy    628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
Db   2250 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2309

Qy    648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
Db   2310 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2369

Qy    667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
Db   2370 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2429

Qy    687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
Db   2430 ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2489

Qy    707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
```

```
Db  2490 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2549

Qy   727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
         |||:::|||||   ||||||||   ||||||   |||   ||||||::::::|||||
Db  2550 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2609

Qy   747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
               ||||||||:::   |||:::|||||||||     |||:::|||     :::|||
Db  2610 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2669

Qy   767 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
         |||||||||||||||||||||::::::|||   |||:::|||:::|||   |||::::::
Db  2670 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 2729

Qy   787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
         |||:::   |||::::::|||         ::::::|||   |||   |||||||||:::
Db  2730 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 2789

Qy   807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
         ||||||:::   ||||||||||||||||||||||||:::||||||   ||||||:::|||
Db  2790 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 2849

Qy   827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
         |||||||||:::|||:::   ||||||   ||||||   :::   ||||||      :::
Db  2850 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2909

Qy   847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
         ||||||:::   :::||||||   |||   |||      ::::::|||:::   |||
Db  2910 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 2969

Qy   867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
         :::||||||:::||||||         ||||||||||||||||||||||:::
Db  2970 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 3029

Qy   887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
         |||:::   |||   :::   :::||||||:::|||:::   :::||||||||||||:::
Db  3030 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 3089

Qy   907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
         |||:::            ||||||:::|||:::||||||||      |||||||||:::
Db  3090 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3149

Qy   927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
            ::::::||||||   |||   |||||||||   |||   ||||||         |||
Db  3150 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3209

Qy   947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
         :::|||   |||||||||:::   |||   ||||||   |||   ||||||   ||||||
Db  3210 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3269

Qy   967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
         |||:::   :::|||:::||||||||||||   :::::::::|||:::|||:::|||
Db  3270 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3329

Qy   987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
         ||||||||||||||||||||||||||||   ||||||||||||:::|||   ||||||
Db  3330 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3389

Qy  1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
         |||:::   |||:::||||||   ||||||   :::   |||   |||||||||||||||
Db  3390 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3449

Qy  1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
         ||||||:::   :::   ||||||   ||||||||||:::||||||:::||||||||||||
Db  3450 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3509

Qy  1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
         ||||||||||||||||||||||||||||||||||||         :::||||||   |||
Db  3510 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3569

Qy  1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
         |||   ||||||||||||:::|||||||||   ||||||   ||||||:::|||:::
Db  3570 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3629

Qy  1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
         ||||||   ||||||   ||||||||||||||||||||||||:::||||||||||||
Db  3630 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3689

Qy  1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
         ||||||:::::::::|||   ||||||||||||||||||||||||||||||||||
Db  3690 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 3749

Qy  1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
         :::||||||||||||||||||   ||||||||||||||||||   ||||||||||||:::
Db  3750 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 3809

Qy  1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
         |||:::   ||||||||||||   ||||||||||||||||||||||||||||      |||
Db  3810 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 3869

Qy  1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
         |||::::::||||||   ||||||||||||::::::||||||      :::   |||||||||
Db  3870 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 3929

Qy  1187 LysLeuValAsnAlaGlnSer 1193
            :::|||:::   |||:::
Db  3930 TCAATGGTCAGTGTCCAGGCT 3950
```

RESULT 14
AAZ94738
ID AAZ94738 standard; cDNA; 4646 BP.
XX
AC AAZ94738;
XX
DT 01-AUG-2000 (first entry)
XX
DE Human ATP binding cassette ABCB1 (MDR1) cDNA.
XX
KW ABCB1; ATP binding cassette; human; cholesterol; lipid disorder;
KW atherosclerosis; lipid disorder; dyslipidemia; psoriasis;
KW lupus erythematosus; diagnosis; gene therapy; MDR1;
KW multidrug resistance; chromosome 7q21; ss.
XX
OS Homo sapiens.
XX
PN WO200018912-A2.
XX
PD 06-APR-2000.
XX
PF 21-SEP-1999; 99WO-EP06991.
XX
PR 25-SEP-1998; 98US-0101706.
XX
PA (FARB) BAYER AG.
XX
PI Schmitz G, Klucken J;
XX
DR WPI; 2000-293151/25.
XX
PT Adenosine triphosphate binding proteins useful for identifying agents
PT for treating atherosclerosis and other inflammatory disorders -
XX
PS Claim 9; Page 110-112; 154pp; English.
XX
CC The present sequence is that of human ATP binding cassette
CC subfamily B protein ABCB1 cDNA. The cDNA was identified using a
CC differential display method in which monocytes from peripheral
CC blood were subjected to macrophage differentiation and cholesterol
CC loading with acetylated low density lipoproteins and subsequent
CC deloading with high density lipoprotein (HDL3) to identify
CC cholesterol sensitive genes. The gene maps to chromosome 7q21
CC and is also termed MDR1 (multidrug resistance). The invention
CC provides cholesterol-sensitive ABC genes (see AAZ94734-63). These
CC genes, and polypeptides encoded by them, can be used for diagnostic
CC and therapeutic applications, and for biochemical or cell-based
CC assays to screen for pharmacologically active modulator compounds

```
CC   useful for the treatment of lipid disorders, atherosclerosis or
CC   other inflammatory diseases such as psoriasis and lupus
CC   erythematosus.
XX
SQ   Sequence 4646 BP; 1371 A; 892 C; 1129 G; 1254 T; 0 other;

Alignment Scores:
Pred. No.:              0            Length:         4646
Score:                  3545.00      Matches:        681
Percent Similarity:     75.55%       Conservative:   246
Best Local Similarity:  55.50%       Mismatches:     264
Query Match:            58.97%       Indels:         36
DB:                     21           Gaps:           9

US-09-873-409-6 (1-1195) x AAZ94738 (1-4646)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    575 ATGGTGGTGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy     21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db    635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db    695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    875 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    935 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 994

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db   1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db   1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db   1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db   1235 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1294

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db   1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1354

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db   1355 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1408

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db   1409 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1468

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db   1469 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db   1529 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1588

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db   1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1648

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   1649 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1708

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   1709 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1768

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   1769 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1828

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   1829 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1888

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1889 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1948

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1949 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2008

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   2009 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2068

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   2069 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2128

Qy    501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db   2129 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 2188

Qy    521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db   2189 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 2248

Qy    541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db   2249 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 2308

Qy    561 Gln-----------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
Db   2309 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 2368

Qy    575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 589
Db   2369 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2428

Qy    590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
Db   2429 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2488

Qy    608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627
```

```
Db   2489 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2548

Qy    628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
Db   2549 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2608

Qy    648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
Db   2609 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2668

Qy    667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
Db   2669 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2728

Qy    687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
Db   2729 ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2788

Qy    707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
Db   2789 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2848

Qy    727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
Db   2849 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2908

Qy    747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
Db   2909 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2968

Qy    767 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
Db   2969 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 3028

Qy    787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
Db   3029 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 3088

Qy    807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
Db   3089 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 3148

Qy    827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
Db   3149 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 3208

Qy    847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
Db   3209 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 3268

Qy    867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
Db   3269 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 3328

Qy    887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
Db   3329 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 3388

Qy    907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
Db   3389 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3448

Qy    927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
Db   3449 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3508

Qy    947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
Db   3509 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3568

Qy    967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
Db   3569 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3628

Qy    987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
Db   3629 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3688

Qy   1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
Db   3689 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3748

Qy   1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
Db   3749 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3808

Qy   1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
Db   3809 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3868

Qy   1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
Db   3869 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3928

Qy   1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
Db   3929 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3988

Qy   1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
Db   3989 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 4048

Qy   1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
Db   4049 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 4108

Qy   1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
Db   4109 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 4168

Qy   1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
Db   4169 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 4228

Qy   1187 LysLeuValAsnAlaGlnSer 1193
Db   4229 TCAATGGTCAGTGTCCAGGCT 4249
```

```
RESULT 15
AAD38994
ID   AAD38994 standard; DNA; 4646 BP.
XX
AC   AAD38994;
XX
DT   23-SEP-2002  (first entry)
XX
DE   Human mdr1 gene.
XX
KW   Human; haematologic malignancy; multidrug resistance; MDR; SUMO-1;
KW   hypoxia inducible factor-1; small ubiquitin-like-modifier; HIF-1;
KW   lymphoid disorder; chronic lymphoproliferative disorder; lymphoma;
KW   myeloid disorder; lymphocytic leukaemia; thrombocythaemia; myeloma;
KW   angiogenic myeloid metaplasia; myeloid leukaemia; gene therapy;
KW   polycythaemia vera; hypoxia responsive element; HRE; gene; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             425..4267
FT                   /*tag=  a
FT                   /product= "Human MDR protein"
XX
PN   WO200234291-A2.
XX
PD   02-MAY-2002.
XX
```

```
PF   25-OCT-2001; 2001WO-US49856.
XX
PR   26-OCT-2000; 2000US-243542P.
XX
PA   (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI   Colgan SP;
XX
DR   WPI; 2002-471427/50.
DR   P-PSDB; AAE24211.
XX
PT   Treating a subject (at risk of) having a hematologic malignancy or
PT   multidrug resistance, e.g. lymphoma or myeloma, by administering
PT   hypoxia inducible factor 1 binding molecules or small
PT   ubiquitin-like-modifier-1 binding molecules  -
XX
PS   Disclosure; Page 53-55; 92pp; English.
XX
CC   The invention relates to a method of treating a subject having or at
CC   risk of developing a haematologic malignancy or multidrug resistance
CC   (MDR). The method involves administering hypoxia inducible factor-1
CC   (HIF-1) binding molecules or small ubiquitin-like-modifier (SUMO)-1
CC   binding molecules or HIF-1-SUMO-1 complex modulators. mdr1-hypoxia
CC   responsive element (HRE) binding molecules or antisense nucleic
CC   acid molecules and SUMO-1 binding molecules or antisense molecules
CC   are useful for treating a subject having or at risk of developing
CC   haematologic malignancy or MDR (e.g. a lymphoid disorder or a myeloid
CC   disorder). The lymphoid disorders include lymphocytic leukaemia or
CC   chronic lymphoproliferative disorders e.g. lymphoma, myeloma or chronic
CC   lymphoid leukaemia. The myeloid disorders include chronic or acute
CC   myeloid leukaemia, e.g. angiogenic myeloid metaplasia, essential
CC   thrombocythaemia or polycythaemia vera. The invention is used in gene
CC   therapy. The present sequence is human mdr1 gene.
XX
SQ   Sequence 4646 BP; 1371 A; 892 C; 1129 G; 1254 T; 0 other;

Alignment Scores:
Pred. No.:               0           Length:        4646
Score:                   3545.00     Matches:       681
Percent Similarity:      75.55%      Conservative:  246
Best Local Similarity:   55.50%      Mismatches:    264
Query Match:             58.97%      Indels:        36
DB:                      24          Gaps:          9

US-09-873-409-6 (1-1195) x AAD38994 (1-4646)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
             |||:::::|||   |||||::::::::::::||||||   |||||||||||    |||
Db     575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy      21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
             |||   ||||||||:::|||                           :::|||:::|||
Db     635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                     ::::::       |||         ||||||
Db     695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                   ||     ||||||   ||||||||         |||:::       |||
Db     755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
             ||||||:::|||   |||   :::   |||   ||||||      :::||||||||||||
Db     815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
             ||||||||:::::::::   |||:::||||||||||||||||   |||:::||||||
Db     875 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
             ||||||||||||:::|||   |||:::   ||||||::::::   ||||||||||||||
Db     935 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 994

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
                :::   ||||||:::|||:::||||||      |||      ||||||      :::
Db     995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
             ||||||||||||||||||||   |||:::   ||||||::::::      |||||||||
Db    1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
                   ::::::::::::   |||   |||   |||||||||   ||||||:::||||||
Db    1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
             ||||||||||||||||||||||||::::::|||||||||||||||||||      |||
Db    1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy     201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
             :::|||||||||:::||||||   :::||||||:::::::||||||      |||||||||
Db    1235 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1294

Qy     221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
             :::   |||   :::   :::||||:::||||||   :::      :::      :::|||
Db    1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1354

Qy     241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
             ||||||||||||||||||||||:::|||:::|||:::||||||      |||:::|||
Db    1355 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1408

Qy     261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
             |||   ||||||   ||||||||||||||||:::      ::::::   :::|||   |||
Db    1409 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1468

Qy     281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
             |||         |||   ||||||      |||||||||         ||||||::::::
Db    1469 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528

Qy     301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
             |||||   ||||||||||||||:::::::|||   :::|||:::|||||||::::::|||
Db    1529 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1588

Qy     321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
             :::|||   :::||||||:::||||||   |||:::|||||||||||||      :::|||
Db    1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1648

Qy     341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
             ||||||||||||||||||||||||:::::::|||||:::||||||||||||||||||
Db    1649 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1708

Qy     361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
             :::|||   |||||||||   ||||||:::||||||||||||||||||      :::
Db    1709 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1768

Qy     381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
             |||   :::   |||||         ||||||||      :::||||||||      |||:::
Db    1769 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1828

Qy     401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
             ||||||||||||||||||||||||||||||   |||||||||:::   ||||||
Db    1829 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1888

Qy     421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
             :::||||||||||:::::::||||||   :::|||:::|||:::|||   :::|||||||||
Db    1889 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1948

Qy     441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
             ||||||||||||||||:::      |||:::|||||||:::|||||||||||||||:::|||
Db    1949 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2008

Qy     461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
             ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db    2009 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2068
```

```
Qy     481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db    2069 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2128

Qy     501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db    2129 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 2188

Qy     521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db    2189 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 2248

Qy     541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db    2249 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 2308

Qy     561 Gln------------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
Db    2309 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 2368

Qy     575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 589
Db    2369 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2428

Qy     590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
Db    2429 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2488

Qy     608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627
Db    2489 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2548

Qy     628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
Db    2549 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2608

Qy     648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
Db    2609 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2668

Qy     667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
Db    2669 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2728

Qy     687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
Db    2729 ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2788

Qy     707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
Db    2789 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2848

Qy     727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
Db    2849 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2908

Qy     747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
Db    2909 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2968

Qy     767 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
Db    2969 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 3028

Qy     787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
Db    3029 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 3088

Qy     807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
Db    3089 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 3148

Qy     827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
Db    3149 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 3208

Qy     847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
Db    3209 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 3268

Qy     867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
Db    3269 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 3328

Qy     887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
Db    3329 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 3388

Qy     907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
Db    3389 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3448

Qy     927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
Db    3449 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3508

Qy     947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
Db    3509 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3568

Qy     967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
Db    3569 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3628

Qy     987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
Db    3629 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3688

Qy    1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
Db    3689 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3748

Qy    1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
Db    3749 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3808

Qy    1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
Db    3809 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3868

Qy    1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
Db    3869 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3928

Qy    1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
Db    3929 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3988

Qy    1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
Db    3989 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 4048

Qy    1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
Db    4049 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 4108

Qy    1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
Db    4109 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 4168

Qy    1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
Db    4169 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 4228

Qy    1187 LysLeuValAsnAlaGlnSer 1193
```

```
Db    4229 TCAATGGTCAGTGTCCAGGCT 4249

Search completed: March 30, 2003, 03:20:36
Job time : 1405.91 secs
```

GenCore version 5.1.4_p5_4578

OM protein - nucleic search, using frame_plus_p2n model

```
Run on:          March 30, 2003, 03:22:44 ; Search time 8219.42 Seconds
                                             (without alignments)
                                             2354.618 Million cell updates/sec

Title:           US-09-873-409-6
Perfect score:   6012
Sequence:        1 MILGILASLVNGACLPLMPL..........QELLRNRDIYFKLVNAQSVQ 1195

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters:      32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115420_19240/app_query.fasta_1.7544
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_30544_@runat_27032003_115420_19240 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       EST:*
                 1:  em_estba:*
                 2:  em_esthum:*
                 3:  em_estin:*
                 4:  em_estmu:*
                 5:  em_estov:*
                 6:  em_estpl:*
                 7:  em_estro:*
                 8:  em_htc:*
                 9:  gb_est1:*
                 10:  gb_est2:*
                 11:  gb_htc:*
                 12:  gb_est3:*
                 13:  gb_est4:*
                 14:  gb_est5:*
                 15:  em_estfun:*
                 16:  em_estom:*
                 17:  gb_gss:*
                 18:  em_gss_hum:*
                 19:  em_gss_inv:*
                 20:  em_gss_pln:*
                 21:  em_gss_vrt:*
                 22:  em_gss_fun:*
                 23:  em_gss_mam:*
                 24:  em_gss_mus:*
                 25:  em_gss_other:*
                 26:  em_gss_pro:*
                 27:  em_gss_rod:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1092	18.2	669	13	BM013981	BM013981 603639619
	2	1051	17.5	1033	14	BM926413	BM926413 AGENCOURT
	3	1008	16.8	943	9	AL520322	AL520322 AL520322
	4	995	16.6	760	13	BM016204	BM016204 603642659
	5	926.5	15.4	929	14	BQ882401	BQ882401 AGENCOURT
	6	918	15.3	545	9	AL040762	AL040762 DKFZp434C
	7	856.5	14.2	948	13	BM471690	BM471690 AGENCOURT
	8	854	14.2	2676	11	AK014319	AK014319 Mus muscu
	9	834	13.9	998	14	BM904842	BM904842 AGENCOURT
	10	771	12.8	894	12	BF584668	BF584668 602098406
	11	770.5	12.8	1019	12	BG248052	BG248052 602359987
	12	753	12.5	726	12	BG293345	BG293345 602390738
	13	752.5	12.5	871	17	AZ682350	AZ682350 ENTKB16TF
	14	751	12.5	944	12	BF796582	BF796582 602258463
	15	744.5	12.4	872	11	AK020318	AK020318 Mus muscu
	16	743.5	12.4	886	17	BH139685	BH139685 ENTNA47TR
c	17	734.5	12.2	932	17	AZ670821	AZ670821 ENTJN69TF
	18	728.5	12.1	947	17	AZ683753	AZ683753 ENTIL96TF
	19	726.5	12.1	886	17	AZ540627	AZ540627 ENTEQ18TF
c	20	723.5	12.0	880	17	AZ687805	AZ687805 ENTLU52TF
	21	721.5	12.0	939	14	BQ720763	BQ720763 AGENCOURT
	22	720.5	12.0	897	17	AZ541090	AZ541090 ENTDS67TR
	23	719	12.0	780	12	BG587938	BG587938 EST489713
	24	711.5	11.8	913	17	BH155700	BH155700 ENTRO54TR
	25	709.5	11.8	834	17	AZ548312	AZ548312 ENTFG07TR
c	26	708.5	11.8	853	17	AZ679807	AZ679807 ENTHI67TR
	27	705.5	11.7	823	17	AZ532602	AZ532602 ENTCR50TR
	28	702.5	11.7	1813	11	U66688	U66688 Homo sapien
	29	697.5	11.6	861	14	BQ717101	BQ717101 AGENCOURT
	30	692.5	11.5	1341	11	AY108285	AY108285 Zea mays
	31	690.5	11.5	899	17	BH154857	BH154857 ENTRO47TF
c	32	690	11.5	835	13	BJ349604	BJ349604 BJ349604
c	33	689.5	11.5	891	17	AZ682250	AZ682250 ENTKT68TR
	34	688	11.4	1123	11	AY108485	AY108485 Zea mays
	35	687	11.4	715	14	BQ869512	BQ869512 QGD6G09.y
c	36	684	11.4	820	13	BJ356540	BJ356540 BJ356540
	37	683	11.4	600	13	BM486593	BM486593 pgm2n.pk0
	38	682.5	11.4	919	17	AZ690701	AZ690701 ENTIY68TR
	39	680	11.3	926	14	BQ123477	BQ123477 EST609053
	40	677.5	11.3	795	12	BF313560	BF313560 601900192
	41	674.5	11.2	815	13	BI409057	BI409057 602961127
	42	674	11.2	2564	11	AK020022	AK020022 Mus muscu
	43	671.5	11.2	803	12	BG298756	BG298756 602396681
	44	670	11.1	852	17	BH720383	BH720383 BOHZZ94TF
	45	669	11.1	547	13	BJ030711	BJ030711 BJ030711

ALIGNMENTS

```
RESULT 1
BM013981
LOCUS       BM013981              669 bp    mRNA    linear   EST 30-OCT-2001
DEFINITION  603639619F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5415556 5',
            mRNA sequence.
ACCESSION   BM013981
VERSION     BM013981.1  GI:16528335
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 669)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
```

```
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12060  row: f  column: 05
            High quality sequence stop: 664.
FEATURES             Location/Qualifiers
     source          1. .669
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5415556"
                     /clone_lib="NIH_MGC_87"
                     /tissue_type="mammary adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.383 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      203 a    122 c    162 g    182 t
ORIGIN

Alignment Scores:
Pred. No.:              6.69e-109     Length:        669
Score:                  1092.00       Matches:       215
Percent Similarity:     99.54%        Conservative:  2
Best Local Similarity:  98.62%        Mismatches:    0
Query Match:            18.16%        Indels:        1
DB:                     13            Gaps:          0

US-09-873-409-6 (1-1195) x BM013981 (1-669)

Qy    265 AlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPhe 284
          :::|||||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db     17 AGTGTTTCTTTAATGTAATCCATAGCAGTTATTGCATTGGAGCAGCAGTCCCTCACTTT 76

Qy    285 GluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLys 304
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     77 GAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTTATTGATAAGAAA 136

Qy    305 ProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrVal 324
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    137 CCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTG 196

Qy    325 GluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGly 344
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    197 GAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGT 256

Qy    345 LeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGly 364
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    257 CTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGG 316

Qy    365 LysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMet 384
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    317 AAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATG 376

Qy    385 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 404
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    377 GTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTG 436

Qy    405 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 424
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    437 GTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGA 496

Qy    425 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 444
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    497 GATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTT 556
```

```
Qy    445 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 464
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    557 ATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATGAGT 616

Qy    465 GlyGlyGlnLysGlnArg-IleAlaIleAlaArgAlaLeuValArgAsnPro 481
          |||||||||||||||||| |||||||||||||||||||||||||||||||||
Db    617 GGAGGGCAGAAACAGAGGCATCGCAATTGCTCGTGCCTTAGTTCGAAACCCC 668

RESULT 2
BM926413
LOCUS       BM926413                1033 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6600788 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5764845
            5', mRNA sequence.
ACCESSION   BM926413
VERSION     BM926413.1  GI:19376792
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1033)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12818  row: o  column: 22
            High quality sequence start: 3
            High quality sequence stop: 686.
FEATURES             Location/Qualifiers
     source          1. .1033
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5764845"
                     /clone_lib="NIH_MGC_114"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
                     male brains, age range 23-27 yo. Library is oligo-dT
                     primed and directionally cloned (EcoRV site is  destroyed
                     upon cloning).  Average insert size 1.5 kb, insert size
                     range 1-3 kb.  Library is normalized and enriched for
                     full-length  clones and was constructed by C. Gruber
                     (Invitrogen).  Research  Genetics tracking code 019. Note:
                     this is a NIH_MGC Library."
BASE COUNT      289 a    212 c    218 g    312 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              4.68e-104     Length:        1033
Score:                  1051.00       Matches:       221
Percent Similarity:     94.49%        Conservative:  2
Best Local Similarity:  93.64%        Mismatches:    13
Query Match:            17.48%        Indels:        3
DB:                     14            Gaps:          0

US-09-873-409-6 (1-1195) x BM926413 (1-1033)

Qy    266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 GTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAA 239

Qy    286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    240 ACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTCATTGATAAGAAACCC 299
```

```
Qy    306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
Db    300 AGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGAA 359

Qy    326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
Db    360 TTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCTG 419

Qy    346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
Db    420 AATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAG 479

Qy    366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
Db    480 AGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGTG 539

Qy    386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
Db    540 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTGGTT 599

Qy    406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
Db    600 AGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGAGAT 659

Qy    426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
Db    660 GATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATC 719

Qy    446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465
Db    720 ATGGAGTTTCCTAATAA-TTTAATACATTGGTAGGGGAANAAGGAGCTCCAATGAGTGGA 778

Qy    466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
Db    779 AGCCAGAACCG-AGGATCCCAATTGCTCCTGCCTTAGTTCGAAACCCCAGGAT-CTGAAT 836

Qy    486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 501
Db    837 TTAGATGAAGCTACGTCTGCCCTGGATTCAGAAACCAGTCAACTGGTC 884

RESULT 3
AL520322
LOCUS       AL520322                943 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL520322 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB006YC15 5
            prime, mRNA sequence.
ACCESSION   AL520322
VERSION     AL520322.1  GI:12783815
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 943)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1..943
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DB006YC15"
                     /clone_lib="LTI_NFL004_NBC2"
                     /sex="male"
                     /tissue_type="neuroblastoma cells"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector.  Library was normalized.  Library was constructed
                     by Life Technologies.  Contact : Feng Liang Life
                     Technologies, a division of Invitrogen 9800 Medical Center
                     Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
                     8371 Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      253 a    225 c    253 g    211 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              2.08e-99        Length:         943
Score:                  1008.00         Matches:        194
Percent Similarity:     80.40%          Conservative:   48
Best Local Similarity:  64.45%          Mismatches:     58
Query Match:            16.77%          Indels:         1
DB:                     9               Gaps:           0

US-09-873-409-6 (1-1195) x AL520322 (1-943)

Qy    894 IleValPheThrAlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAla 913
Db      5 TTATTATTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGCAAGTCAGTTCATTTGCT 64

Qy    914 ProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLys 933
Db     65 CCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACC 124

Qy    934 ProAsnIleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeu 953
Db    125 CCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTC 184

Qy    954 GluPheArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGly 973
Db    185 ACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGA 244

Qy    974 LeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGly 993
Db    245 CTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGTGGG 304

Qy    994 LysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeu 1013
Db    305 AAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTGCTG 364

Qy   1014 PheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIle 1033
Db    365 CTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGCATC 424

Qy   1034 ValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAsp 1053
Db    425 GTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGAGAC 484

Qy   1054 AsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHis 1073
Db    485 AACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATACAT 544

Qy   1074 SerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGln 1093
Db    545 GCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACTCAG 604

Qy   1094 LeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLys 1113
Db    605 CTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCTCAT 664

Qy   1114 IleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGln 1133
Db    665 ATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTCCAA 724

Qy   1134 HisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSer 1153
Db    725 GAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATTGTGATTGCTCACCGCCTGTCC 784

Qy   1154 AlaIleGlnAsnAlaAspLeuIleVal-ValLeuHisAsnGlyLysIleLysGluGlnGl 1173
```

```
              ||||||||||||||||||||||| |||      ||||||::::::||||||    ||
Db     785 ACCATCCAGAATGCAGACTTAATAGTGTGTTTCAGAATGGCAGAGTCAAGGAGCATGG 844

Qy    1173 yThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSe 1193
           |||||||||:::||||||      :::   |||||||||   :::|||:::    |||::
Db     845 CACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTTTCAATGGTCAGTGTCCAGGC 904

Qy    1193 r 1193
           :
Db     905 T 905

RESULT 4
BM016204
LOCUS       BM016204                 760 bp    mRNA    linear   EST 30-OCT-2001
DEFINITION  603642659F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5418615 5',
            mRNA sequence.
ACCESSION   BM016204
VERSION     BM016204.1  GI:16530558
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 760)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12068  row: e  column: 16
            High quality sequence stop: 740.
FEATURES             Location/Qualifiers
     source          1..760
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5418615"
                     /clone_lib="NIH_MGC_87"
                     /tissue_type="mammary adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.383 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      229 a    150 c    187 g    193 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              3.76e-98      Length:         760
Score:                  995.00        Matches:        203
Percent Similarity:     98.07%        Conservative:   0
Best Local Similarity:  98.07%        Mismatches:     2
Query Match:            16.55%        Indels:         2
DB:                     13            Gaps:           0

US-09-873-409-6 (1-1195) x BM016204 (1-760)

Qy     303 LysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGly 322
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3 AAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGA 62

Qy     323 ThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeu 342
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63 ACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTG 122

Qy     343 LysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGly 362
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 AAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGC 182

Qy     363 SerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPhe 382
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 AGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTT 242

Qy     383 IleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle 402
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 ATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATN 302

Qy     403 GlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr 422
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 GGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTAT 362

Qy     423 GlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyr 442
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     363 GGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTAT 422

Qy     443 AspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGln 462
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     423 GATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAA 482

Qy     463 MetSerGlyGlyGlnLysGlnArgIleAlaIle-AlaArgAlaLeuValArgAsnProLy 482
           |||||||||||||||||||||||||||||||||   ||||||||||||||||||||||||
Db     483 ATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGGCTCGTGCCTTAGTTCGAAACCCCAA 542

Qy     482 sIleLeuIleLeu-AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValG 502
           ||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db     543 GATTCTGATTTTCAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTC 602

Qy     502 lnAlaAlaLeuGluLys 507
           ||||||||||||||||
Db     603 AAGCTGCACTGGAGAAG 619

RESULT 5
BQ882401
LOCUS       BQ882401                 929 bp    mRNA    linear   EST 16-AUG-2002
DEFINITION  AGENCOURT_8627902 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6291782
            5', mRNA sequence.
ACCESSION   BQ882401
VERSION     BQ882401.1  GI:22274409
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 929)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2492  row: c  column: 15
            High quality sequence stop: 677.
FEATURES             Location/Qualifiers
     source          1..929
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6291782"
                     /clone_lib="NIH_MGC_43"
                     /tissue_type="normal pigmented retinal epithelium"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming. Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
```

```
                 adaptor: GGCACGAG(G).  Library constructed  by Ling Hong
                 in the laboratory of Gerald M. Rubin (University of
                 California, Berkeley) using ZAP-cDNA synthesis kit
                 (Stratagene) and  Superscript II RT (Life Technologies).
                 Note: this is a NIH_MGC Library. |"
BASE COUNT      281 a    176 c    237 g    234 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:               1.8e-90       Length:        929
Score:                   926.50        Matches:       188
Percent Similarity:      79.30%        Conservative:  61
Best Local Similarity:   59.87%        Mismatches:    58
Query Match:             15.41%        Indels:        7
DB:                      14            Gaps:          2

US-09-873-409-6 (1-1195) x BQ882401 (1-929)

Qy     176 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThr-ValIl 195
           ||||||:::|||||||||||||||||||||||||||||||::::::|||||||| |||||
Db       1 GCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGGTGAT 60

Qy     195 eAlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLy 215
           ||||||       |||:::||||||||:::||||||   :::||||||::::::||||
Db      61 TGCATTTGGAGGACAAAAGAAAGAACTTGAAAGGTACAACAAAAAATTTAGAAGAAGCTAA 120

Qy     215 sAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePh 235
           |     ||||||||:::    |||    :::    :::|||:::||||||    :::
Db     121 AAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCT 180

Qy     235 eMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGl 255
            :::        :::|||  |||||||||||||||||||||:::|||:::|||:::|||||
Db     181 GATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGA 240

Qy     255 uProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTy 275
           |       |||:::||||||   ||||||   |||||||||||||||:::         :::::
Db     241 A------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTT 294

Qy     275 rCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPh 295
           :    :::|||   |||   |||         |||   ||||||   ||||||||||||||::
Db     295 TAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTA 354

Qy     295 eHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTy 315
           :      ||||||::::::||||||     ||||||||||||||::::::|||     :::|||::
Db     355 TGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCA 414

Qy     315 rLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSe 335
           :||||||:::::|||:::|||        :::||||||:::||||||     |||:::|||||||
Db     415 CAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATC 474

Qy     335 rArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrVa 355
           ||||      :::||||||||||||||||||||||||:::::::::||||||:::|||||
Db     475 TCGAAAAGAAGTTAAGATCTTGAAGGGTCTGAACCTGAAGGTGCAGAGTGGGCAGACGGT 534

Qy     355 lAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLe 375
           ||||||||||||      :::|||    ||||||||||||     |||||||||:::|||||||
Db     535 GGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCT 594

Qy     375 uTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVa 395
           |||||||||      :::|||    :::        ||||||         ||||||||     :::|||||
Db     595 CTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGT 654

Qy     395 lArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrTh 415
           ||||          |||:::   ||||||||||||||||||||||||||||||        |||||
Db     655 AAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCAC 714

Qy     415 rIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAl 435
           ||||:::    ||||||    ||||||||::::::||||||      :::|||:::|||:::||
Db     715 GATAGCTGAAAACATTCCCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGC 774

Qy     435 aAlaArgGluAlaAsnAla-TyrAspPheIleMetGluPheProAsnLysPheAsnThrL 455
           |     :::|||||||||||| ||||||||||||||:::     |||:::||||||:::||||
Db     775 TGTCCAGGAAGCCAATGCCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCC 834

Qy     455 euValGlyGluLysGlyAlaGlnMetSerGlyGly---GlnLysGlnArgIleAlaIleA 474
           |||||||     :::       |||:::||||||||     |||     :::          :::|
Db     835 TGGTTGGANAAAGAAGGGGCCAATTGAGTGGTGGGGCACAAACCAAAGGATCTCCCTTGG 894

Qy     474 laArg-AlaLeuVal-ArgAsnProLysIleLeu 484
           ||||| |||||||||   ||||||||||||||
Db     895 CACGTGGCCCTGGTTCCCCAACCCCAAGATCCTC 928

RESULT 6
AL040762
LOCUS       AL040762             545 bp    mRNA    linear   EST 29-FEB-2000
DEFINITION  DKFZp434C1815_r1 434 (synonym: htes3) Homo sapiens cDNA clone
            DKFZp434C1815 5', mRNA sequence.
ACCESSION   AL040762
VERSION     AL040762.1  GI:5409708
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 545)
  AUTHORS   Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
            Wiemann,S.
  TITLE     EST (Bloecker, et al.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Bloecker H
            MIPS
            Am Klopferspitz 18a D-82152 Martinsried, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by GBF (National Research Centre for Biotechnology Ltd.,
            Braunschweig/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            s1 sequence also available.
            This clone (DKFZp434C1815) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1..545
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp434C1815"
                     /clone_lib="434 (synonym: htes3)"
                     /tissue_type="testis"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT      171 a     89 c    138 g    147 t
ORIGIN

Alignment Scores:
Pred. No.:               5.84e-90      Length:        545
Score:                   918.00        Matches:       181
Percent Similarity:      100.00%       Conservative:  0
Best Local Similarity:   100.00%       Mismatches:    0
Query Match:             15.27%        Indels:        0
DB:                      9             Gaps:          0

US-09-873-409-6 (1-1195) x AL040762 (1-545)

Qy     306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3 AGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGAA 62

Qy     326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63 TTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCTG 122

Qy     346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 AATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAG 182
```

```
Qy     366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 AGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGTG 242

Qy     386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTGGTT 302

Qy     406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 AGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGAGAT 362

Qy     426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     363 GATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATC 422

Qy     446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     423 ATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGA 482

Qy     466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     483 GGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATT 542

Qy     486 Leu 486
           |||
Db     543 TTA 545
```

RESULT 7
BM471690

```
LOCUS       BM471690                 948 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6465349 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5539117
            5', mRNA sequence.
ACCESSION   BM471690
VERSION     BM471690.1  GI:18520732
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 948)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12233  row: b  column: 14
            High quality sequence stop: 569.
FEATURES             Location/Qualifiers
     source          1..948
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5539117"
                     /clone_lib="NIH_MGC_72"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.  Library constructed by Life
                     Technologies."
BASE COUNT      285 a    187 c    194 g    282 t
ORIGIN

Alignment Scores:
Pred. No.:              9.03e-83      Length:         948
Score:                     856.50     Matches:        186
Percent Similarity:        89.67%     Conservative:   5
Best Local Similarity:     87.32%     Mismatches:     14
Query Match:               14.25%     Indels:         8
DB:                        13         Gaps:           1

US-09-873-409-6 (1-1195) x BM471690 (1-948)

Qy     266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     205 GTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAA 264

Qy     286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     265 ACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTTATTGATAAGAAACCC 324

Qy     306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     325 AGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGAA 384

Qy     326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     385 TTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCTG 444

Qy     346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     445 AATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAG 504

Qy     366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     505 AGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGTG 564

Qy     386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle-GlyValVa 405
           |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db     565 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTTGGAGTGGT 624

Qy     405 lSerGlnGluPro--ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr-GlyAr 424
           |||||||||||||  |||||||||||||||||||||||||||||||||||||||| |||||
Db     625 TAGTCAAGAGCCCTGGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGGACG 684

Qy     424 gAspAspValThrAspGluGlu-MetGluArgAlaAlaArgGluAla-AsnAlaTyrAsp 443
           ||||||     |||||||||::: ||||||:::|||   ||||||       |||    :::
Db     685 AGATGATGGGACTGATGAACAAAATGGAAAAAGCAACAAGGGAAACCAAAGGCGAATGAA 744

Qy     444 PheIleMetGlu-PheProAsnLysPheAsnThrLeuVal---GlyGluLysGlyAlaGl 462
           ||||||||||| ||||||  |||||||||||||||       |||:::|||       ||
Db     745 TTTATCATGGAATTTCCCAAAAAAATTTAATACATTGGAAAGGGGAAAAAAGGACCTTCA 804

Qy     462 nMetSerGlyGlyGlnLysGlnArg 470
           |         |||    |||    :::
Db     805 AGGAATGGAAGGGCCAAAACCCAAG 829
```

RESULT 8
AK014319

```
LOCUS       AK014319                2676 bp    mRNA    linear   HTC 19-JAN-2002
DEFINITION  Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length
            enriched library, clone:3222401P09:ATP-binding cassette, sub-family
            B (MDR/TAP), member 8, full insert sequence.
ACCESSION   AK014319
VERSION     AK014319.1  GI:12852089
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (strain:C57BL/6J) 14, 17 days embryo head cDNA to
            mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:3222401P09.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
```

```
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409 (6821), 685-690 (2001)
  MEDLINE   21085660
   PUBMED   11217851
REFERENCE   5  (bases 1 to 2676)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C.,
            Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,
            Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F.,
            Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,
            Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,
            Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
            Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,
            Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and
            Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Gemomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. Second
            strand cDNA was prepared with the primer adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cleaved
            with BamHI and XhoI. cDNA of size comprised between 3 and 7 kb was
            selected before cloning. Vector: a modified pBluescript KS(+) after
            bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'
            end: BamHI. Host: DH10B.
FEATURES             Location/Qualifiers
     source          1..2676
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:3222401P09"
                     /db_xref="MGD:MGI:1907488"
                     /db_xref="taxon:10090"
                     /clone="3222401P09"
                     /tissue_type="head"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="14, 17 days embryo"
     gene            1..2676
                     /gene="Abcb8"
     CDS             72..2225
                     /gene="Abcb8"
                     /note="ATP-binding cassette, sub-family B (MDR/TAP),
                     member 8
                     data source:MGD, source key:MGI:1351667, evidence:ISS
                     putative"
                     /codon_start=1
                     /protein_id="BAB29270.1"
                     /db_xref="GI:12852090"
                     /db_xref="MGD:MGI:1351667"
                     /translation="MLVHLFRFGIRGGPVPGWSLQSLRFQTFSAARSSDDRLSSHLLR
                     TVAQLRVQLRAHLPRAPPASHWSPSAWCWVGGTLVVPAVLWQHPRLCLIALCEAKESP
                     PAQPTRAPELRFNWKLFWHFLHPHLLALGAAIVLALGAALVNVQIPLLLGQLVEIVAK
                     YTRDHMGSFVSESRKLSVQLLLLYGVQGLLTFGYLVLLSHIGERMAMDMRKALFSSLL
                     RQDIAFFDAKKTGQLVSRLTTDVQEFKSSFKLVISQGLRSCTQVIGSLVSLSMLSPRL
                     TLMLAVVTPALMGVGTLMGSGLRKLSRQCQEQIARATGVADEALGNVRTVRAFAMEKR
                     EERYQAELESCCCKAEELGRGIALFQGLSNIAFNCMVLGTLFIGGSLVAGQQLKGGD
                     LMSFLVASQTVQRSMASLSVLFGQVVRGLSAGARVFEYMALSPVIPLTGGYCIPNKDI
                     RGSITFQNVTFSYPCRPGFNVLKDFTLKLPSGKIVALVGQSGGGKTTVASLLERFYDP
                     EAGSVTLDGHDLRTLNPSWLRGQVIGFISQEPVLFATTIMENIRFGKLDASDEEVYTA
                     AREANAHEFISSFPDGYSTVVGERGTTLSGGQKQRLAIARALIKQPTVLILDEATSAL
                     DAESERVVQEALDRASAGRTVLVIAHRLSTVRAAHSIIVMANGQVCEAGTHEELLKKG
                     GLYSELIRRQTLDASLTSTPPAEKPEDPKSCQSKA"
     polyA_signal    2653..2658
                     /gene="Abcb8"
                     /note="putative"
     polyA_site      2676
                     /gene="Abcb8"
                     /note="putative"
BASE COUNT      544 a    785 c    756 g    591 t
ORIGIN

Alignment Scores:
Pred. No.:              1.12e-81      Length:         2676
Score:                  854.00        Matches:        221
Percent Similarity:     55.74%        Conservative:   109
Best Local Similarity:  37.33%        Mismatches:     202
Query Match:            14.20%        Indels:         60
DB:                     11            Gaps:           19

US-09-873-409-6 (1-1195) x AK014319 (1-2676)

Qy          1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
                ::::::|||   :::   |||:::|||||||||                :::||||||
Db        471 ATTGTGCTGGCCTTAGGTGCGGCACTAGTGAAT------------GTGCAGATCCCCTTG 518

Qy         21 ValLeuGlyGluMetSerAspAsnLeuIleSer------------------GlyCysLeu 34
                :::||||||::::::   :::   :::                         |||
Db        519 CTCCTGGGCCAGCTGGTGGAGATTGTCGCCAAGTACACGAGGGACCACATGGGGAGTTTC 578
```

```
Qy     35 ValGlnThrAsnThrTyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyVal 54
Db    579 GTGTCTGAGTCCCGTAAGCTCAGCGTCCAGCTGCTCCTACTGTAC---------GGTGTT 629

Qy     55 AlaAlaLeuIle---PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArg 73
Db    630 CAGGGACTGCTGACCTTTGGATACCTAGTGCTGCTGTCCCACATT------GGTGAGCGC 683

Qy     74 GlnThrLysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrp 93
Db    684 ATGGCCATGGACATGCGGAAAGCCCTTTTCAGCTCCCTGCTCCGGCAAGACATTGCTTTC 743

Qy     94 PheAspSerCysAspIleGlyGluLeuAsnThrArgMet---ThrAspIleAspLys--- 111
Db    744 TTTGATGCCAAAAAGACAGGGCAGCTAGTGAGTCGCTTGACTACTGATGTGCAAGAATTC 803

Qy    112 ---------------------IleSerAspGlyIleGlyAspLysIleAlaLeuLeuPhe 124
Db    804 AAGTCATCCTTCAAGCTTGTCATCTCCCAGGGACTG------------------------ 839

Qy    125 GlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu---ValLysGlyTrpLys 143
Db    840 ---CGCAGCTGCACCCAGGTGATTGGTAGCCTGGTGTCCCTGTCTATGCTGTCCCCTCGC 896

Qy    144 LeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSer 163
Db    897 CTTACCCTGATGCTGGCTGTCGTCACACCCGCCCTCATGGGAGTGGGCACCCTGATGGGC 956

Qy    164 ArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaVal 183
Db    957 TCAGGCCTCCGAAAGCTGTCTCGCCAGTGTCAGGAGCAGATTGCCAGGGCAACAGGTGTA 1016

Qy    184 AlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGlu 203
Db   1017 GCAGATGAGGCCCTTGGCAATGTTCGGACTGTGCGGGCCTTCGCGATGGAGAAGAGGGAG 1076

Qy    204 LeuGlnArgTyrThrGlnAsnLeuLysAsp------------AlaLysAspPheGlyIle 219
Db   1077 GAGGAACGCTATCAAGCAGAACTGGAGTCATGCTGCTGTAAAGCAGAAGAACTGGGC--- 1133

Qy    220 LysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhe------PheMetAsn 237
Db   1134 ---AGGGGCATCGCCTTGTTCCAAGGGCTCTCCAACATCGCTTTCAACTGTATGGTCTTG 1190

Qy    238 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGlu---Pro 256
Db   1191 GGCACC------------CTGTTCATTGGGGGCTCCCTTGTGGCTGGACAGCAGCTGAAA 1238

Qy    257 GlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCys 276
Db   1239 GGGGGAGACCTCATGTCCTTCCTGGTGGCTTCCCAGACAGTACAGAGGTCTATGGCCAGC 1298

Qy    277 IleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGly------AlaAla 294
Db   1299 CTCTCTGTCCTGTTTGGTCAG---------------GTGGTACGTGGGCTGAGTGCGGGA 1343

Qy    295 PheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGly 314
Db   1344 GCCCGAGTCTTCGAATACATGGCCCTGAGCCCTGTCATCCCA------TTGACCGGGGGC 1397

Qy    315 Tyr------LysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsn 332
Db   1398 TACTGCATCCCCAACAAGGACATTCGTGGTTCCATCACCTTCCAAAATGTCACCTTCAGC 1457

Qy    333 TyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGly 352
Db   1458 TACCCCTGCAGACCTGGCTTCAATGTGCTCAAGGACTTCACCCTGAAGCTGCCCTCTGGC 1517

Qy    353 GluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeu 372
Db   1518 AAGATTGTGGCTCTTGTGGGCCAGTCTGGGGGAGGAAAGACCACAGTTGCCTCCTTGCTG 1577

Qy    373 GlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAla 392
Db   1578 GAACGCTTCTATGACCCTGAAGCTGGCTCGGTGACGTTGGATGGCATGACCTGCGAACT 1637

Qy    393 LeuAsnValArgHisTyrArgAspHis---IleGlyValValSerGlnGluProValLeu 411
Db   1638 CTCAACCCCTCCTGGCTCCGGGGCCAGGTCATAGGTTTCATCAGCCAGGAGCCAGTCCTG 1697

Qy    412 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 431
Db   1698 TTTGCAACAACCATCATGGAGAATATCCGATTTGGGAAGCTGGATGCTTCCGATGAAGAG 1757

Qy    432 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 451
Db   1758 GTGTACACAGCTGCACGAGAAGCCAATGCCCACGAGTTCATCAGCAGCTTCCCCGATGGC 1817

Qy    452 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 471
Db   1818 TACAGCACTGTGGTTGGTGAGCGGGGCACAACCTTGTCTGGTGGCCAGAAGCAGCGCCTA 1877

Qy    472 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 491
Db   1878 GCCATCGCACGTGCCCTCATCAAGCAGCCCACAGTGCTGATCCTGGACGAGGCCACCAGT 1937

Qy    492 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAlaSerLysGly 511
Db   1938 GCGCTAGATGCAGAATCCGAGAGGGTGGTACAGGAGGCCCTGGACCGGGCCAGTGCTGGC 1997

Qy    512 ArgThrThrIleValValAlaHisArgLeuSerThrIleArgSerAlaAspLeuIleVal 531
Db   1998 CGCACCGTGTTGGTCATTGCCCACCGGCTTAGTACTGTCCGTGCAGCCCACTCCATCATT 2057

Qy    532 ThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGluLeuMetAlaLysArg 551
Db   2058 GTCATGGCCAATGGCCAAGTCTGTGAGGCTGGGACCCACGAAGAACTCCTTAAAAAAGGC 2117

Qy    552 GlyLeuTyrTyrSerLeuValMetSerGlnAspIle 563
Db   2118 GGGCTCTATTCAGAGCTTATCCGGAGACAAACCCTG 2153

RESULT 9
BM904842
LOCUS       BM904842                 998 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6699581 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5557655
            5', mRNA sequence.
ACCESSION   BM904842
VERSION     BM904842.1  GI:19355221
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 998)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12279  row: f  column: 24
            High quality sequence stop: 738.
FEATURES             Location/Qualifiers
     source          1..998
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5557655"
                     /clone_lib="NIH_MGC_72"
```

```
                 /tissue_type="melanotic melanoma"
                 /lab_host="DH10B (phage-resistant)"
                 /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                 Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                 Average insert size 2 kb.   Library constructed by Life
                 Technologies."
BASE COUNT      271 a    223 c    248 g    256 t
ORIGIN

Alignment Scores:
Pred. No.:              2.93e-80        Length:        998
Score:                  834.00          Matches:       164
Percent Similarity:     74.12%          Conservative:  68
Best Local Similarity:  52.40%          Mismatches:    77
Query Match:            13.87%          Indels:        4
DB:                     14              Gaps:          0

US-09-873-409-6 (1-1195) x BM904842 (1-998)

Qy    718 AspIleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAla 737
            ||:::::||||||||:::   :::|||:::|||||  ||||||||  ||||||
Db     61 GATGTGAGTTGGTTTGATGACCCTAAAAACACCACTGGAGCATTGACTACCAGGCTCGCC 120

Qy    738 IleAspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsn 757
             |||     ||||||::::::|||||    |||||||::    |||:::|||||||||
Db    121 AATGATGCTGCTCAAGTTAAAGGGGCTATAGGTTCCAGGCTTGCTGTAATTACCCAGAAT 180

Qy    758 AlaThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPhe 777
                |||:::|||        :::||||||||||||||||||||||||::::::|||
Db    181 ATAGCAAATCTTGGGACAGGAATAATTATATCCTTCATCTATGGTTGGCAACTAACACTG 240

Qy    778 LeuIleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMet 797
            |||:::|||:::|||   |||::::::|||:::   |||::::::|||         :::
Db    241 TTACTCTTAGCAATTGTACCCATCATTGCAATAGCAGGAGTTGTTGAAATGAAAATGTTG 300

Qy    798 ThrGlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGlu 817
            :::|||    |||   |||||||||:::||||||:::    ||||||||||||||||||
Db    301 TCTGGACAAGCACTGAAAGATAAGAAAGAACTAGAAGGTGCTGGGAAGATCGCTACTGAA 360

Qy    818 AlaLeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMet 837
            |||:::||||||   ||||||:::|||||||||||||:::|||:::    ||||||  |||
Db    361 GCAATAGAAAACTTCCGAACCGTTGTTTCTTTGACTCAGGAGCAGAAGTTTGAACATATG 420

Qy    838 TyrGluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGly 857
            |||   :::   ||||||       :::||||||:::    :::||||||   |||   |||
Db    421 TATGCTCAGAGTTTGCAGGTACCATACAGAAACTCTTTGAGGAAAGCACACATCTTTGGA 480

Qy    858 SerCysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPhe 877
                  ::::::|||:::   |||   :::||||||:::||||||     |||||||||
Db    481 ATTACATTTTCCTTCACCCAGGCAATGATGTATTTTTCCTATGCTGGATGTTTCCGGTTT 540

Qy    878 GlyAlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThr 897
            ||||||||||||:::              |||:::   |||    :::    :::||||||:::
Db    541 GGAGCCTACTTGGTGGCACATAAACTCATGAGCTTTGAGGATGTTCTGTTAGTATTTTCA 600

Qy    898 AlaIleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSer 917
            |||:::   :::|||||||||||||||:::|||:::             ||||||:::|||:::
Db    601 GCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTTGCTCCTGACTATGCC 660

Qy    918 LysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAsp 937
            |||||||||      ||||||||||:::         ::::::||||||   |||   ||||||
Db    661 AAAGCCAAAATATCAGCAGCCCACATCATCATGATCATTGAAAAAACCCCTTTGATTGAC 720

Qy    938 SerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGlu 957
            |||   |||   ||||||      ||||:::|||   |||||||||:::   |||
Db    721 AGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAATGTCACATTTGGGTGA 780

Qy    958 ValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer-LeuSe 977
            |||   |||   ||||||   ||||||||:::     :::|||:::|||||||||  |||
Db    781 GTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAGGGACTGAGCCCTGGA 840

Qy    977 rIleGluArgGlyLysThrValAlaPheVal-GlySerSerGlyCysGlyLysSerThrS 997
             ::::::::|||:::|||:::|||   |||  ||||||||||||||||||||||||||
Db    841 GGTGAAGAAGGGGCAGACCCTGGCTCTGGTGGGGCAGCAGTGGCTGTGGGAAGAGCACAG 900

Qy    997 er-ValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGly 1016
               ||||||||||||:::|||:::|||:::|||:::       ::::::|||   ||||||
Db    901 TGGGTCCAGCTCCTGGAGCGGGTCTACAACCCCTTGGCAAGGAAAATGCTGCTTGATGGC 960

Qy   1017 ValAspAlaLys-GluLeuAsnValGlnTrp 1026
                      |||    |||:::|||   |||
Db    961 AAAAGAATAAAACCGACTGGATGTTCCATGG 991

RESULT 10
BF584668
LOCUS       BF584668                  894 bp    mRNA    linear   EST 12-DEC-2000
DEFINITION  602098406F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4218385 5',
            mRNA sequence.
ACCESSION   BF584668
VERSION     BF584668.1  GI:11658386
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 894)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9798  row: 1  column: 02
            High quality sequence stop: 651.
FEATURES             Location/Qualifiers
     source          1..894
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4218385"
                     /clone_lib="NCI_CGAP_Co24"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.6 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      263 a    203 c    242 g    185 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              1.98e-73        Length:        894
Score:                  771.00          Matches:       171
Percent Similarity:     79.72%          Conservative:  57
Best Local Similarity:  59.79%          Mismatches:    48
Query Match:            12.82%          Indels:        15
DB:                     12              Gaps:          3

US-09-873-409-6 (1-1195) x BF584668 (1-894)

Qy    266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285
            ||||||||||||||:::          ::::::   :::|||   |||   |||:::   |||
Db      5 GTCTTCTTTTCCGTGTTAAT-GGAGCATTCAGTGTTGGACAGGCATCTCCAAATATTGAA 63

Qy    286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305
               ||||||   ||||||||||||||||:::   :::|||::::::||||||   ||||||
Db     64 GCCTTCGCCAATGCACGAGGAGCAGCTTATGAAGTCTTCAAAATAATTGATAATAAGCCC 123

Qy    306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
            ||||||||:::||||||    :::|||:::||||||::::::|||:::|||    :::|||
Db    124 AGTATAGACAGCTTCTCAAAGAGTGGGCACAAACCAGACAACATACAAGGAAATCTGGAA 183
```

```
Qy      326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
            |||||||||:::     |||:::||||||||||||      ::::::||||||||||||||
Db      184 TTTAAGAATATTCACTTCAGTTACCCATCTCGAAAAGAAGTTCAGATCTTGAAGGGCCTC 243

Qy      346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
            ||||||::::::|||||||||:::|||||||||||||||||||     :::|||   ||||||
Db      244 AATCTGAAGGTGAAGAGCGGACAGACGGTGGCCCTGGTTGGCAACAGTGGCTGTGGAAAA 303

Qy      366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
            ||||||   |||||||||:::||||||||||||||||||     ||||||    :::   :::
Db      304 AGCACAACTGTCCAGCTGATGCAAAGGCTCTACGACCCCCTAGATGGCATGGTCAGTATC 363

Qy      386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
            |||         |||||||||   :::|||||||||:::   |||:::   ||||||||||||
Db      364 GACGGACAGGACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCATTGGTGTGGTG 423

Qy      406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
            ||||||||||||||||||||||||      ||||||||||:::   ||||||:::||||||||:::
Db      424 AGTCAGGAACCTGTGCTGTTTGCCACCACGATCGCCGAGAACATTCGCTATGGCCGAGAA 483

Qy      426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
            |||||||||   :::|||:::|||:::|||   :::||||||||||||||||||||||||||||
Db      484 GATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTCATC 543

Qy      446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465
            |||:::   |||::::::|||:::|||||||||||||||||||:::||||||||||:::||||||
Db      544 ATGAAACTGCCCCACCAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGG 603

Qy      466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
            ||||||||||||||||||   ||||||||||||||||||||||||||||||||   |||
Db      604 GGACAGAAACAGAGAATCGC-ATTGCCCGGGCCCTGGTCCGCAATCCCAAGTGCCTTTC- 661

Qy      486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla-Le 505
            :::||||||||||||||||||   ||||||:::||||||::::::      |||||||||||| ||
Db      662 GTGGACGAAGCCACCTCAGC-CTGGATACAGAAAGTGAAGCTGGGGTTCAGGCCGCACCT 720

Qy      505 uGluLysAlaSerLysGlyArgThrThrIle------------ValValAlaHisArgLe 521
            |::::::         ||||||   ||||||:::           ::::::     :::||
Db      721 AGATAGGGTACAAAAGGGCNGGACCACATTGGGAAGCTCATCGCTGTTACCGTCGAAGCT 780

Qy      521 uSerThrIleArgSerAlaAspLeuIleVal------ThrLeuLysAspGlyMetLeuAl 539
            |:::   :::|||            :::|||      |||||||||               ::
Db      781 GACGGCCATGCGGCCGAAGGGGGG-CTTGTGGGAACGACTCTGAAAAGC---------TC 830

Qy      539 aGluLysGlyAlaHis 544
            :|||||||||   |||
Db      831 TGAAAAAGGTACTCAT 846
```

```
RESULT 11
BG248052
LOCUS       BG248052                1019 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  602359987F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:4488404 5',
            mRNA sequence.
ACCESSION   BG248052
VERSION     BG248052.1  GI:12757867
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 1019)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Gilbert Smith, Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10334  row: n  column: 21
            High quality sequence stop: 650.
FEATURES             Location/Qualifiers
     source          1..1019
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4488404"
                     /clone_lib="NCI_CGAP_Mam1"
                     /tissue_type="tumor, biopsy sample"
                     /dev_stage="10 months, virgin"
                     /lab_host="DH10B"
                     /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
                     Site_2: NotI; Cloned unidirectionally.  Primer: Oligo dT.
                     Library constructed by Life Technologies. Investigator
                     providing samples: Gilbert Smith, NIH"
BASE COUNT      276 a    262 c    283 g    196 t     2 others
ORIGIN

Alignment Scores:
Pred. No.:              2.85e-73     Length:         1019
Score:                  770.50       Matches:        154
Percent Similarity:     79.17%       Conservative:   36
Best Local Similarity:  64.17%       Mismatches:     47
Query Match:            12.82%       Indels:         3
DB:                     12           Gaps:           1

US-09-873-409-6 (1-1195) x BG248052 (1-1019)

Qy      950 GluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgProAspValPhe 969
            |||||||||:::::::|||         |||    |||   ||||||   ||||||      :::
Db        4 GAAGGAAATGTGCAATTTAGTGGAGTCGTGTTCAACTATCCCACCCGACCCAGCATCCCA 63

Qy      970 IleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPheValGlySer 989
            :::|||:::||||||||||||      :::::::::|||:::|||:::|||   |||||||||
Db       64 GTGCTTCAGGGGCTGAGCCTTGAGGTGAAGAAGGGCCAGACGCTGGCCCTGGTGGGCAGC 123

Qy      990 SerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGln 1009
            ||||||||||||||||||||||   ||||||||||||:::|||   |||||||||:::
Db      124 AGTGGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAGCGCTTCTACGACCCCATGGCT 183

Qy     1010 GlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSer 1029
            |||   |||         ||||||      :::      |||:::|||||||||||||||||||:::
Db      184 GGATCAGTGTTTCTAGATGGCAAAGAAATAAAGCAACTGAATGTCCAGTGGCTCCGAGCA 243

Qy     1030 GlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAlaGluAsnIle 1049
            |||:::      ||||||   |||||||||:::||||||:::||||||||||||||||||||||
Db      244 CAGCTGGGCATTGTGTCCCAAGAGCCCATTCTCTTTGACTGCAGCATCGCAGAGAACATT 303

Qy     1050 AlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAlaAlaAsnAla 1069
            |||||||||||||||||||||||||||         :::||||||      ||||||
Db      304 GCCTACGGAGACAACAGCCGGGTCGTGTCTTATGAGGAGATTGTGAGGGCAGCCAAGGAG 363

Qy     1070 AlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeu 1089
            ||||||||||||   ||||||:::   ||||||:::|||||||||||||||:::||||||
Db      364 GCCAACATCCACCAGTTCATCGACTCGCTACCTGATAAATACAACACCAGAGTAGGAGAC 423

Qy     1090 LysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeu 1109
            ||||||   ||||||||||||||||||||||||||||:::||||||||||||||||||:::
Db      424 AAAGGCACTCAGCTGTCGGGTGGGCAGAAGCAGCGCATCGCCATCGCACGCGCCCTCGTC 483

Qy     1110 GlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsnAspSerGlu 1129
            ::::::|||   |||||||||||||||||||||||||||||||||||||     :::||||||
Db      484 AGACAGCCTCACATTTTACTTCTGGACGAAGCAACATCAGCTCTGGATACAGAAAGTGAA 543

Qy     1130 LysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThr 1149
            ||||||||||||   ||||||||||||||||||||   ||||||||||||:::|||:::
Db      544 AAGGTTGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACCTGCATTGTGATCGCT 603

Qy     1150 HisArgLeuSerAlaIleGlnAsnAlaAspLeuIle-ValValLeuHis-AsnGlyLysI 1169
            ||||||               |||||||||||||||||| |||||||:::     |||   |||:
Db      604 CACCGCTGTCCACCATCCCAGAACGCGGACTTGATCGGTGGTGATTCAAGAACCGGAAGG 663
```

```
Qy    1169 leLys---GluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
             ::|||      :::          |||:::::::::       :::   ||||||||
Db     664 TCAAGGGAGCACGGGCACCCCAACAGCAAGTGGTGGCGCAGAAGGGCATCTACTTC 719

RESULT 12
BG293345
LOCUS       BG293345                726 bp    mRNA    linear   EST 21-FEB-2001
DEFINITION  602390738F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4502552 5',
            mRNA sequence.
ACCESSION   BG293345
VERSION     BG293345.1  GI:13052943
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 726)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10371  row: 1  column: 09
            High quality sequence stop: 658.
FEATURES             Location/Qualifiers
     source          1..726
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4502552"
                     /clone_lib="NIH_MGC_94"
                     /tissue_type="retina"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 3.3 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      189 a    220 c    190 g    127 t
ORIGIN

Alignment Scores:
Pred. No.:              1.28e-71      Length:        726
Score:                  753.00        Matches:       152
Percent Similarity:     84.29%        Conservative:  25
Best Local Similarity:  72.38%        Mismatches:    30
Query Match:            12.52%        Indels:        3
DB:                     12            Gaps:          0

US-09-873-409-6 (1-1195) x BG293345 (1-726)

Qy     984 ValAlaPheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArg 1003
             :::|||   ||||||||||||||||||||||||||||||     ||||||||||||:::|||
Db       2 CTGGCCCTGGTGGGCAGCAGTGGCTGCGGGAAGAGCACAGTGGTCCAGCTGCTCGAGCGC 61

Qy    1004 LeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsn 1023
              |||||||||:::   |||    ||||||    ||||||    :::||||||:::||||||
Db      62 TTCTATGACCCCATGGCTGGATCAGTGCTCTTAGATGGTCAAGAAGCAAAGAAACTCAAT 121

Qy    1024 ValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCys 1043
             ||||||||||||||:::|||:::      ||||||   |||||||||:::||||||:::|||
Db     122 GTCCAGTGGCTCCGAGCTCAACTGGGCATTGTGTCCCAGGAACCCATTCTCTTTGACTGC 181

Qy    1044 SerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIle 1063
             ||||||||||||||||||||||||||||||||||||||||||||||||     |||||||||
Db     182 AGCATCGCAGAGAACATCGCCTATGGAGACAACAGCCGGGTCGTGCCTCATGATGAGATT 241

Qy    1064 LysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyr 1083
                ||||||     ||||||||||||     |||||||||     ||||||:::||||||
Db     242 GTGAGGGCAGCCAAGGAGGCCAACATCCACCCCTTCATCGAGACGCTGCCCCAAAAATAT 301

Qy    1084 AsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAla 1103
             ||||||:::||||||     ||||||     ||||||||||||||||||||||||||||:::|||
Db     302 AACACAAGAGTAGGAGACAAGGGGACGCAGCTCTCTGGGGGCCAGAAGCAGAGGATTGCC 361

Qy    1104 IleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAla 1123
             ||||||||||||||||:::::::::|||::::::||||||||||||||||||||||||||||
Db     362 ATCGCCCGAGCCCTCATCAGACAGCCTCGGGTCCTACTGCTGGATGAAGCCACGTCAGCT 421

Qy    1124 LeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArg 1143
             ||||||    :::||||||||||||||||||     ||||||||||||||||||     ||||||
Db     422 CTGGATACTGAGAGTGAAAAGGTTGTCCAGGAAGCACTGGACAAAGCCAGGGAAGGCCGC 481

Qy    1144 ThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla-AspLeuIleValVa 1163
             ||||||:::    :::   |||||||||||||     ||||||||||||     ||||||||||
Db     482 ACCTGCATTGCGATCGCTCACCGCCTGTCCACCATCCAGAACGCGGCACTTGATCGTGGT 541

Qy    1163 lLeuHisAsnGlyLysIleLysGluGlnGlyThr-HisGlnGluLeuLeuArgAsnArgA 1183
             |:::    ||||||||:::||||||     ||||||  ||||||:::||||||       :::
Db     542 GATTGAGAACGGCAAGGTCAAGGAGCACGGCACCCCACCAGCAGCTGCTGGCGCAGAAGG 601

Qy    1183 spIleTyrPhe-LysLeuValAsn 1190
                ||||||||| :::|||||||||
Db     602 GCATCTATTTCTCAACTGGTCAAC 625

RESULT 13
AZ682350
LOCUS       AZ682350                871 bp    DNA     linear   GSS 14-DEC-2000
DEFINITION  ENTKB16TF Entamoeba histolytica Sheared DNA Entamoeba histolytica
            genomic, DNA sequence.
ACCESSION   AZ682350
VERSION     AZ682350.1  GI:11819496
KEYWORDS    GSS.
SOURCE      Entamoeba histolytica.
  ORGANISM  Entamoeba histolytica
            Eukaryota; Entamoebidae; Entamoeba.
REFERENCE   1  (bases 1 to 871)
  AUTHORS   Loftus,B., Van Aken,S. and Fraser,C.
  TITLE     Determination of clone end sequences from Entamoeba histolytica
            HM1:IMSS sheared DNA library
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Brendan J Loftus
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0208
            Fax: 301 838 3543
            Email: bjloftus@tigr.org
            Clones are derived from the Entamoeba histolytica HM1:IMSS sheared
            DNA library
            Seq primer: M13-Forward
            Class: shotgun
            High quality sequence start: 16
            High quality sequence stop: 860.
FEATURES             Location/Qualifiers
     source          1..871
                     /organism="Entamoeba histolytica"
                     /strain="HM1:IMSS"
                     /db_xref="taxon:5759"
                     /clone_lib="Entamoeba histolytica Sheared DNA"
                     /note="Vector: pHOS1; Site_1: Bst I; Constructed at The
                     Institute for Genomic Research (TIGR), Rockville, MD.
                     Genomic DNA isolated from broth cultures of E. histolytica
                     using a  method described by Clark and Diamond (Clark,
                     C.G., and Diamond, L.S.  (1993) Entamoeba histolytica: a
                     method for isolate identification.  Exp. Parasitol.
                     77:450.). The DNA was mechanically sheared to give a
                     tight size distribution (~2 kb). The v + i method used for
                     the library  construction is described in detail in Smith,
```

```
              H.O. and Venter, J.C.  (Making small insert libraries for
              whole genome shotgun sequencing  projects. In Genome
              Sequencing: A Practical Approach, eds. M. Vaudin  and B.
              Barell, Oxford University Press, 1999)."
BASE COUNT      349 a    117 c    174 g    231 t
ORIGIN

Alignment Scores:
Pred. No.:              2.03e-71        Length:         871
Score:                  752.50          Matches:        151
Percent Similarity:     71.64%          Conservative:   46
Best Local Similarity:  54.91%          Mismatches:     75
Query Match:            12.52%          Indels:         3
DB:                     17              Gaps:           1

US-09-873-409-6 (1-1195) x AZ682350 (1-871)

Qy    290 AlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsn 309
          |||:::   |||||||||::::::::|||   ||||||:::   |||   |||||
Db     45 GCTAAAGTTGCCGCATTTAATGTTTATCAAACAATTGATAGAATCCCAGATATTGATTGT 104

Qy    310 PheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnVal 329
             |||      |||      |||                |||   :::   |||::::::|||
Db    105 CAGTCTATTGGAGGTGAATGTCCAACTGAGTGTAATGGAAATATTAGATTTGAGGATGTT 164

Qy    330 SerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIle 349
             |||   ||||||:::|||   |||         :::||||||||||||:::|||   |||
Db    165 CAATTTGTTTATCCAACAAGACTGTCTCATCATGTATTAAAAGGACTTGACCTTGAAATT 224

Qy    350 LysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValVal 369
          |||   |||:::|||:::||||||||||||   :::|||   ||||||||||||   :::
Db    225 AAGAAAGGACAAACAATTGCATTAGTTGGAGCATCAGGATGTGGGAAGTCAACTACTATT 284

Qy    370 GlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAsp 389
          ||||||:::||||||||||||:::   |||   :::   :::|||         |||
Db    285 CAATTAATCCAAAGAAATTATGATCCAAATGGTGGAAGAGTAACATTAGACGGAAAAGAT 344

Qy    390 IleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluPro 409
          ||||||   ||||||::::::         |||:::   ||||||:::|||   |||||||||
Db    345 ATACGAGAGTTGAATATCAAATGGTTAAGAAATCAAATAGGATTAGTTGGACAAGAACCA 404

Qy    410 ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAsp--------- 426
          |||||||||      ||||||      ||||||      |||      :::
Db    405 GTGTTGTTTGCAGGAACAATTCGAGAAAATATTATGCTTGGAGCTAAAGAAGGAGCAACA 464

Qy    427 ValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMet 446
             ::::::|||||||||         |||:::   |||||||||:::|||||||||
Db    465 CCAAGTGAAGAAGAGATGATTGAATGTGCTAAAATGGCAAATGCACATGACTTCATTTCT 524

Qy    447 GluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGly 466
          :::   |||         ::::::|||::::::|||||||||||||||      :::||||||||||
Db    525 AAACTTCCAGAAGGATATGACACAATAATTGGAGAAAAAGGAGCATTATTATCAGGAGGA 584

Qy    467 GlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeu 486
          ||||||||||||||||||||||||||||||:::||||||||   ||||||:::|||
Db    585 CAAAAACAAAGAATTGCAATTGCACGTGCATTGATTCGAAACCCATCTATTCTTCTTCTT 644

Qy    487 AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGlu 506
          ||||||||||||||||||||||||:::::::|||:::      ||||||   |||||||||
Db    645 GATGAAGCTACATCAGCACTTGATACACAAAGTGAAAAGATTGTACAAGAAGCACTTGAA 704

Qy    507 LysAlaSerLysGlyArgThrThrIleValValAlaHisArgLeuSerThrIleArgSer 526
          |||||||||||||||||||||||||||:::||||||||||||||||:::|||:::|||:::
Db    705 AAAGCATCTAAAGGAAGAACAACAATTATTGTAGCACATAGACTAACAACTGTTAGAAAT 764

Qy    527 AlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGluLysGlyAlaHisAlaGlu 546
          ||||||   |||               |||   :::   |||:::|||   |||   |||
Db    765 GCAGATAAAATATGTGTATTTCATCAAGGAGAAATTATTGAACAAGGAAAACATCAAGAA 824

Qy    547 LeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSerGln 561
          ||||||      :::|||   ||||||   ||||||      |||
Db    825 TTAATGGATTTGAAAGGAACATATTATGGATTAGTCAAAAGACAA 869
```

```
RESULT 14
BF796582
LOCUS       BF796582                 944 bp    mRNA    linear   EST 12-JAN-2001
DEFINITION  602258463F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4341710 5',
            mRNA sequence.
ACCESSION   BF796582
VERSION     BF796582.1  GI:12101636
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 944)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Louis Staudt, M.D., Ph.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9955  row: n  column: 15
            High quality sequence start: 9
            High quality sequence stop: 669.
FEATURES             Location/Qualifiers
     source          1..944
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4341710"
                     /clone_lib="NIH_MGC_85"
                     /tissue_type="lymphoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: lymph; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.867 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      265 a    241 c    237 g    201 t
ORIGIN

Alignment Scores:
Pred. No.:              3.43e-71        Length:         944
Score:                  751.00          Matches:        166
Percent Similarity:     74.62%          Conservative:   31
Best Local Similarity:  62.88%          Mismatches:     60
Query Match:            12.49%          Indels:         9
DB:                     12              Gaps:           2

US-09-873-409-6 (1-1195) x BF796582 (1-944)

Qy    916 TyrSerLysAlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsn 935
          |||:::|||||||||         ||||||||||||||||||   |||   |||:::::::|||
Db     39 TATGCTAAAGCTAAGCTGTCTGCAGCCCACTTATTCATGCTGTTTGAAAGACAACCTCTG 98

Qy    936 IleAspSerArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPhe 955
          ||||||||||   |||:::||||||   |||||||||   |||||||||:::   |||
Db     99 ATTGACAGCTACAGTGAAGAGGGGCTGAAGCCTGATAAATTTGAAGGAAATATAACATTT 158

Qy    956 ArgGluValSerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSer 975
             ||||||   |||   ||||||   |||      :::|||   :::|||:::||||||||||
Db    159 AATGAAGTCGTGTTCAACTATCCCACCCGAGCAAACGTGCCAGTGCTTCAGGGGCTGAGC 218

Qy    976 LeuSerIleGluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSer 995
          |||   :::::::::|||:::|||:::|||   ||||||||||||||||||||||||||||
Db    219 CTGGAGGTGAAGAAAGGCCAGACACTAGCCCTGGTGGGCAGCAGTGGCTGTGGGAAGAGC 278

Qy    996 ThrSerValGlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAsp 1015
          |||   ||||||||||||:::|||   ||||||||||:::   |||   ||||||   |||
Db    279 ACGGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCGGGGACAGTGCTTCTCGAT 338
```

```
Qy    1016 GlyValAspAlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValPro 1035
           |||   :::||||||:::||||||||||||||||||||||:::|||:::   ||||||
Db     339 GGTCAAGAAGCAAAGAAACTCAATGTCCAGTGGCTCAGAGCTCAACTCGGAATCGTGTCT 398

Qy    1036 GlnGluProValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSer 1055
           |||||||||:::||||||:::||||||||||||||||||||||||||||||||||||||||
Db     399 CAGGAGCCTATCCTATTTGACTGCAGCATTGCCGAGAATATTGCCTATGGAGACAACAGC 458

Qy    1056 ArgValValProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPhe 1075
           |||||||||      |||||||||      ||||||   |||||||||||||||    |||
Db     459 CGGGTTGTATCACAGGATGAAATCGTGAGTGCAGCCAAAGCTGCCAACATACATCCTTTC 518

Qy    1076 IleGluGlyLeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSer 1095
           ||||||   ||||||   ||||||   |||:::||||||   ||||||   |||||||||
Db     519 ATCGAGACGTTACCCCACAAATATGAAACAAGAGTGGGAGATAAGGGGACTCAGCTCTCA 578

Qy    1096 GlyGlyGlnLysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLys-ProLysIleLe 1115
           ||||||   |||||||||:::||||||||||||||||||:::::::::: |||:::|||||
Db     579 GGAGGTCACAAACAGAGGATTGCTATTGCCCGAGCCCTCATCAGACAACCCTCAAATCCT 638

Qy    1115 uLeuLeuAspGluAlaThrSerAla-LeuAspAsnAspSerGluLysValVal------G 1133
           |||||||||||||||||||||||| ||||||                |||       :
Db     639 CCTGTTGGATGAAGCTACATCAGCCTCTGGATACCTGACAAGTGAACACGGTTGTCCCAA 698

Qy    1133 lnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeuValValThrHisArgLeuS 1153
           ::                |||           ||||||||||||||:::   ||||||
Db     699 GAAGCCCCTGTCACACAGCAGAAGAGCGCCGCACCTGCCTTGTGATTGCTCACCGCTGTC 758

Qy    1153 erAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnG 1173
                 |||||||||||||||  |||||||||           |||::::::|||
Db     759 AAC--ATCCAGAATGCAGACTT-ATAGTGGTGTTCCCAACGGGAGAAGTCAAG------- 808

Qy    1173 lyThrHis 1175
             ||||||
Db     809 --ACACAT 814
```

RESULT 15
AK020318
LOCUS   AK020318   872 bp   mRNA   linear   HTC 19-JAN-2002
DEFINITION Mus musculus adult male epididymis cDNA, RIKEN full-length enriched
            library, clone:9230106F14:similar to MULTIDRUG RESISTANCE PROTEIN 3
            (P-GLYCOPROTEIN 3) (MDR1A), full insert sequence.
ACCESSION   AK020318
VERSION   AK020318.1 GI:12860872
KEYWORDS   HTC; CAP trapper.
SOURCE   Mus musculus (strain:C57BL/6J) adult male epididymis cDNA to mRNA,
            clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:9230106F14.
  ORGANISM Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE   High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE   Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE   RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
  TITLE   Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409 (6821), 685-690 (2001)
  MEDLINE   21085660
   PUBMED   11217851
REFERENCE   5 (bases 1 to 872)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C.,
            Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,
            Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F.,
            Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,
            Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,
            Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
            Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,
            Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and
            Hayashizaki,Y.
  TITLE   Direct Submission
  JOURNAL   Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Gemomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT   Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. cDNA went
            through one round of normalization to Rot = 10.0 and subtraction to
            Rot = 185.2. Second strand cDNA was prepared with the primer
            adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cleaved
            with BamHI and XhoI. Vector: a modified pBluescript KS(+) after
            bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'
            end: BamHI. Host: DH10B.
FEATURES   Location/Qualifiers
     source   1. .872
                     /organism="Mus musculus"

```
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:9230106F14"
                     /db_xref="MGD:MGI:1911565"
                     /db_xref="taxon:10090"
                     /clone="9230106F14"
                     /sex="male"
                     /tissue_type="epididymis"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     CDS             86..>872
                     /note="data source:SPTR, source key:P21447, evidence:ISS
                     putative
                     similar to MULTIDRUG RESISTANCE PROTEIN 3 (P-GLYCOPROTEIN
                     3) (MDR1A)"
                     /codon_start=1
                     /protein_id="BAB32070.1"
                     /db_xref="GI:12860873"
                     /db_xref="MGD:MGI:1924956"
                     /translation="MANSERTNGLQETNQRYGPLQEQVPKVGNQAVGPIEIFRFADNL
                     DIVLMTLGILASMINGATVPLMSLVLGEISDHLINGCLVQTNRTKYQNCSQTQEKLNE
                     DIIVLTLYYIGIGAAALIFGYVQISFWVITAARQTTRIRKQFFHSILAQDISWFDGSD
                     ICELNTRMTGDINKLCDGIGDKIPLMFQNISGFSIGLVISLIKSWKLSLVVLSTSPLI
                     MASSALCSRMIISLTSKELDAYSKAGAVAEEALSSIQTVTAFGA"
BASE COUNT      262 a    178 c    187 g    245 t
ORIGIN

Alignment Scores:
Pred. No.:              1.53e-70       Length:         872
Score:                  744.50         Matches:        151
Percent Similarity:     80.37%         Conservative:   21
Best Local Similarity:  70.56%         Mismatches:     27
Query Match:            12.38%         Indels:         15
DB:                     11             Gaps:           2

US-09-873-409-6 (1-1195) x AK020318 (1-872)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          |||   ||||||||||||||||::::::||||||||    :::||||||||    |||
Db    230 ATGACTCTGGGAATATTAGCATCCATGATAAATGGAGCCACCGTTCCTTTAATGTCCCTG 289

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
          ||||||||||||:::|||||:::||||||:::|||||||||||||||||||
Db    290 GTTTTAGGAGAAATAAGTGATCATTTAATTAATGGATGCCTAGTACAAACTAACAGAACT 349

Qy     41 SerPhe------------------------------------------PheArgLeuThr 46
             :::                                              ||||||
Db    350 AAATATCAGAACTGTTCTCAGACTCAAGAAAAGCTGAATGAAGATATCATTGTGTTGACT 409

Qy     47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
          ||||||||:::||||||||    |||||||||||||||||||:::|||||||||
Db    410 CTATATTATATTGGAATAGGAGCAGCTGCCCTCATTTTTGGCTATGTACAGATTTCCTTC 469

Qy     67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
          |||:::|||||||||||||||||     |||||||||||||||||||||||||:::
Db    470 TGGGTCATAACTGCAGCCCGGCAAACCACAAGAATCCGAAAACAGTTTTTTCATTCAATT 529

Qy     87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
          ||||||||||||||   ||||||||      ||||||   |||||||||||||||||
Db    530 TTGGCACAAGACATCAGCTGGTTTGATGGCAGTGACATCTGTGAACTTAACACCCGCATG 589

Qy    107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
          |||   |||||:::|||:::    ||||||||||||||||||||   |||:::||||||
Db    590 ACTGGTGACATCAACAAACTCTGTGATGGTATTGGAGATAAGATCCCTCTGATGTTTCAG 649

Qy    126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
          |||:::|||   ||||||||||||||    :::    |||:::|||   |||||||||:::
Db    650 AACATATCTGGGTTTTCTATTGGCCTGGTGATAAGTTTGATAAAAAGCTGGAAACTCTCC 709

Qy    146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
          ||||||   |||||||||||||||||||||||||||||||:::|||   ||||||||||||
Db    710 CTGGTGGTTCTGTCTACATCTCCTCTCATAATGGCTTCATCGGCACTGTGCTCTAGGATG 769

Qy    166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
             :::|||||||||||||||||||||||   ||||||||||||||||||||||||||||
Db    770 ATTATCTCATTGACCAGCAAGGAGCTGGATGCCTATTCCAAAGCTGGGGCTGTGGCTGAA 829

Qy    186 GluValLeuSerSerIleArgThrValIleAlaPheArgAla 199
          |||   |||||||||||:::||||||   ||||||   |||
Db    830 GAAGCCTTGTCATCAATCCAAACAGTCACAGCCTTTGGAGCC 871

Search completed: March 31, 2003, 13:51:24
Job time : 8245.55 secs
```

GenCore version 5.1.4_p5_4578


```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 31, 2003, 02:24:20 ; Search time 175.002 Seconds
                                           (without alignments)
                                           2094.145 Million cell updates/sec

Title:          US-09-873-409-6
Perfect score:  6012
Sequence:       1 MILGILASLVNGACLPLMPL..........QELLRNRDIYFKLVNAQSVQ 1195

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters:      882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115422_19259/app_query.fasta_1.7544
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_578_@runat_27032003_115422_19259 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Issued_Patents_NA:*
                1:  /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
                2:  /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
                3:  /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
                4:  /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
                5:  /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
                6:  /cgn2_6/ptodata/2/ina/backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	3545	59.0	4646	1	US-08-181-471-2	Sequence 2, Appli
2	3545	59.0	6505	2	US-08-793-610-5	Sequence 5, Appli
3	3545	59.0	9318	2	US-08-793-610-6	Sequence 6, Appli
4	3543.5	58.9	4264	2	US-08-784-649A-1	Sequence 1, Appli
5	3543.5	58.9	4264	2	US-08-784-649A-5	Sequence 5, Appli
6	3543	58.9	4669	6	5206352-3	Patent No. 5206352
7	3527	58.7	4669	2	US-08-752-447-1	Sequence 1, Appli
8	3527	58.7	4669	4	US-09-316-167-1	Sequence 1, Appli
9	3517.5	58.5	4233	3	US-09-120-513-1	Sequence 1, Appli
10	3517.5	58.5	4233	4	US-09-450-105-1	Sequence 1, Appli
11	3497	58.2	4669	2	US-08-583-276-18	Sequence 18, Appl
12	2282.5	38.0	4002	2	US-08-996-545-1	Sequence 1, Appli
13	2282.5	38.0	4002	2	US-08-996-545-3	Sequence 3, Appli
14	2282.5	38.0	4002	4	US-09-328-320-1	Sequence 1, Appli
15	2282.5	38.0	4002	4	US-09-328-320-3	Sequence 3, Appli
16	2280	37.9	2726	1	US-08-461-823-1	Sequence 1, Appli
17	2273.5	37.8	4047	2	US-08-612-734B-1	Sequence 1, Appli
18	2211	36.8	4800	2	US-08-612-734B-3	Sequence 3, Appli
19	2166	36.0	3924	1	US-08-395-246C-1	Sequence 1, Appli
20	2116	35.2	4224	1	US-08-612-521-1	Sequence 1, Appli
21	1731.5	28.8	6143	1	US-08-612-521-3	Sequence 3, Appli
22	1703	28.3	3924	2	US-08-996-644-3	Sequence 3, Appli
23	1703	28.3	3924	3	US-09-352-552-3	Sequence 3, Appli
24	1703	28.3	3927	2	US-08-996-644-1	Sequence 1, Appli
25	1703	28.3	3927	3	US-09-352-552-1	Sequence 1, Appli
26	1664	27.7	3909	1	US-08-232-537-1	Sequence 1, Appli
27	1659.5	27.6	3792	4	US-09-351-224E-10	Sequence 10, Appl
28	1525	25.4	3999	4	US-09-351-224E-9	Sequence 9, Appli
29	1109	18.4	13188	4	US-08-961-527-70	Sequence 70, Appl
30	1057	17.6	7760	4	US-08-961-527-63	Sequence 63, Appl
31	996.5	16.6	4411529	4	US-09-103-840A-1	Sequence 1, Appli
32	994.5	16.5	4403765	4	US-09-103-840A-2	Sequence 2, Appli
33	945.5	15.7	7186	4	US-08-961-527-39	Sequence 39, Appl
34	938	15.6	6492	4	US-08-961-527-188	Sequence 188, App
35	761.5	12.7	2376	1	US-08-394-880B-1	Sequence 1, Appli
36	718	11.9	1749	4	US-09-134-001C-1893	Sequence 1893, Ap
37	698	11.6	2061	4	US-09-061-764A-17	Sequence 17, Appl
38	687.5	11.4	5889	1	US-08-463-092B-5	Sequence 5, Appli
39	687.5	11.4	5889	2	US-08-462-109A-5	Sequence 5, Appli
40	687.5	11.4	5889	2	US-08-460-907B-5	Sequence 5, Appli
41	687.5	11.4	5889	3	US-08-463-179A-5	Sequence 5, Appli
42	687.5	11.4	5889	3	US-08-461-384B-5	Sequence 5, Appli
43	674.5	11.2	5120	3	US-08-772-270A-6	Sequence 6, Appli
44	674.5	11.2	8370	2	US-08-488-706-1	Sequence 1, Appli
45	671.5	11.2	5011	1	US-08-463-092B-3	Sequence 3, Appli

ALIGNMENTS

```
RESULT 1
US-08-181-471-2
; Sequence 2, Application US/08181471
; Patent No. 5641508
;  GENERAL INFORMATION:
;    APPLICANT:  Li, Lingna
;    APPLICANT:  Lishko, Valeryi K.
;    TITLE OF INVENTION:  METHOD FOR DELIVERING BENEFICIAL
;    TITLE OF INVENTION:  COMPOSITIONS TO HAIR FOLLICLES
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Thomas Fitting
;      STREET:  12526 High Bluff Drive, Suite 300
;      CITY:  San Diego
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  92130
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/181,471
;      FILING DATE:  13-JAN-1994
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/041,553
;      FILING DATE:  02-APR-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fitting, Thomas
;      REGISTRATION NUMBER:  34,163
;      REFERENCE/DOCKET NUMBER:  ANT0029P
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  619-792-3680
;      TELEFAX:  619-792-8477
```

```
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4646 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  425..4267
US-08-181-471-2

Alignment Scores:
Pred. No.:                 0             Length:         4646
Score:                     3545.00       Matches:        681
Percent Similarity:        75.55%        Conservative:   246
Best Local Similarity:     55.50%        Mismatches:     264
Query Match:               58.97%        Indels:         36
DB:                        1             Gaps:           9

US-09-873-409-6 (1-1195) x US-08-181-471-2 (1-4646)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy     21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db    635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db    695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    875 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    935 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 994

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db   1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db   1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db   1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db   1235 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1294

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db   1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1354

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db   1355 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1408

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db   1409 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1468

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db   1469 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db   1529 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1588

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db   1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1648

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   1649 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1708

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   1709 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1768

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   1769 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1828

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   1829 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1888

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1889 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1948

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1949 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2008

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   2009 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2068

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   2069 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2128

Qy    501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db   2129 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 2188

Qy    521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db   2189 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 2248

Qy    541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db   2249 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 2308

Qy    561 Gln-----------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
Db   2309 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 2368

Qy    575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 589
```

```
Db   2369 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2428

Qy    590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
             :::|||::::::   |||         |||:::         :::   ::::::
Db   2429 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2488

Qy    608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627
          :::   |||:::|||   ||||||      :::|||:::||||||||       ||||||
Db   2489 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2548

Qy    628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
          |||:::   |||:::|||         ::::::::||||||   :::   |||    |||
Db   2549 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2608

Qy    648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
          :::|||||||||:::||||||||      :::|||         :::|||      |||
Db   2609 GCAATAATATTTCAAAGATTATAGGGGTTTTACAAGAATTGATGATCCTGAAACAAAA 2668

Qy    667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
          :::   ::::::   ::::::|||::::::|||:::   ||||||:::|||   |||:::
Db   2669 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2728

Qy    687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
          ::::::|||:::||||||      :::|||:::||||||||||||||||||   ||||||
Db   2729 ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2788

Qy    707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
          |||::::::   |||::::::||||||   |||::::::|||||||||:::      :::
Db   2789 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2848

Qy    727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
          |||:::||||||   |||||||||   ||||||   |||   ||||||::::::||||||
Db   2849 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2908

Qy    747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
             |||||||||:::   |||:::|||||||||      |||:::|||      :::|||
Db   2909 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2968

Qy    767 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
          ||||||||||||||||||||||||::::::|||   |||:::|||:::|||   |||::::::
Db   2969 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 3028

Qy    787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
          |||:::   |||::::::|||         ::::::|||   |||   |||||||||:::
Db   3029 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 3088

Qy    807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
          ||||||:::   ||||||||||||||||||||||||||||||:::   ||||||:::|||
Db   3089 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 3148

Qy    827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
          |||||||||:::|||:::   ||||||   ||||||   :::   ||||||      :::
Db   3149 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 3208

Qy    847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
          ||||||:::   :::||||||   |||   |||      ::::::|||:::   |||
Db   3209 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 3268

Qy    867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
          :::||||||:::||||||         ||||||||||||||||||||||:::
Db   3269 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 3328

Qy    887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
          |||:::   |||   :::   :::||||||:::|||:::|||||||||||||||:::
Db   3329 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 3388

Qy    907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
          |||:::            |||||:::|||:::|||||||||      |||||||||:::
Db   3389 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3448

Qy    927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
                ::::::||||||   |||   |||||||||   |||   ||||||      |||
Db   3449 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3508

Qy    947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
          :::|||   |||||||||:::   |||   ||||||   |||   ||||||   ||||||
Db   3509 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3568

Qy    967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
          |||:::   :::|||:::||||||||||||      :::::::::|||:::|||:::|||
Db   3569 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3628

Qy    987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
          |||||||||||||||||||||||||||      ||||||||||||:::|||   ||||||
Db   3629 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3688

Qy   1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
          |||:::   |||:::||||||      ||||||   :::   |||   |||||||||||||||
Db   3689 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3748

Qy   1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
          ||||||:::   :::   ||||||   |||||||||:::||||||:::||||||||||||
Db   3749 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3808

Qy   1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
          ||||||||||||||||||||||||||||||||||||      :::||||||      |||
Db   3809 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3868

Qy   1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
          |||      ||||||||||||:::|||||||||   ||||||   ||||||:::|||:::
Db   3869 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3928

Qy   1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
          ||||||   ||||||      ||||||||||||||||||||||||:::||||||||||||
Db   3929 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3988

Qy   1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
          ||||||:::::::::|||   ||||||||||||||||||||||||||||||||||
Db   3989 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 4048

Qy   1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
          :::||||||||||||||||||   ||||||||||||||||||   ||||||||||||:::
Db   4049 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 4108

Qy   1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
          |||:::   ||||||||||||   ||||||||||||||||||||||||         |||
Db   4109 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 4168

Qy   1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
          |||::::::||||||   ||||||||||||:::||||||      :::   |||||||||
Db   4169 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 4228

Qy   1187 LysLeuValAsnAlaGlnSer 1193
             :::|||:::   |||:::
Db   4229 TCAATGGTCAGTGTCCAGGCT 4249
```

RESULT 2
US-08-793-610-5
; Sequence 5, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
;   APPLICANT: BAUM, Christopher
;   APPLICANT: STOCKING-HARBERS, Carol
;   APPLICANT: OSTERTAG, Wolfram
;   TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEROF
;   TITLE OF INVENTION: FOR GENE TRANSFER
;   NUMBER OF SEQUENCES: 6
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
;   STREET: 655 Fifteenth Street N.W. Suite 330
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: U.S.A.
;   ZIP: 20005-5701
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk

```
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/793,610
;      FILING DATE:  07-MAR-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P 44 31 973.8
;      FILING DATE:  08-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 195 03 952.1
;      FILING DATE:  07-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/03175
;      FILING DATE:  10-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berman, Richard J.
;      REGISTRATION NUMBER:  39,105
;      REFERENCE/DOCKET NUMBER:  P1614-7007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)638-5000
;      TELEFAX:  (202)638-4810
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6505 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA
US-08-793-610-5

Alignment Scores:
Pred. No.:                0              Length:          6505
Score:                    3545.00        Matches:         681
Percent Similarity:       75.55%         Conservative:    246
Best Local Similarity:    55.50%         Mismatches:      264
Query Match:              58.97%         Indels:          36
DB:                       2              Gaps:            9

US-09-873-409-6 (1-1195) x US-08-793-610-5 (1-6505)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            |||:::::|||   |||||:::::::::::|||||   |||||||||||    |||
Db   1967 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 2026

Qy     21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
            |||    ||||||||:::|||                          :::|||:::|||
Db   2027 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 2086

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                           ::::::      |||          ||||||
Db   2087 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 2146

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                |||         ||||||   ||||||||          |||:::      |||
Db   2147 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 2206

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
            ||||||:::|||   |||    :::    |||   ||||||       :::||||||||||||
Db   2207 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 2266

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
            ||||||||:::::::::   |||:::||||||||||||||||      |||:::||||||
Db   2267 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 2326

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
            ||||||||||||:::|||   |||:::   |||||||::::::   |||||||||||||||
Db   2327 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 2386

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
               :::   ||||||:::|||:::||||||          |||        ||||||        :::
Db   2387 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 2446

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
            ||||||||||||||||||||||   |||:::   ||||||::::::       |||||||||
Db   2447 GGTTGGAAGCTAACCCTTGTGATTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 2506

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
               ::::::::::::   |||   |||   ||||||||   ||||||:::||||||
Db   2507 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 2566

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
            |||||||||||||||||||||::::::||||||||||||||||||       |||
Db   2567 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 2626

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
            :::|||||||||:::||||||   :::||||||:::::::||||||      ||||||||
Db   2627 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 2686

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
            :::   |||   :::   :::|||:::||||||    :::       :::      :::|||
Db   2687 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 2746

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
            ||||||||||||||||||||||:::|||:::||||||              |||:::|||
Db   2747 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 2800

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
            |||   ||||||   |||||||||||||||:::       ::::::   :::|||   |||
Db   2801 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 2860

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
               |||      |||   ||||||   ||||||||||||||:::   ||||||::::::
Db   2861 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 2920

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
            ||||||   |||||||||||||||:::::::|||      :::|||:::||||||::::::|||
Db   2921 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 2980

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
            :::|||   :::||||||:::||||||   |||:::||||||||||||       :::|||
Db   2981 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 3040

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
            ||||||||||||||||||||:::::::::||||||:::||||||||||||||||||
Db   3041 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 3100

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
            :::|||   ||||||||||||   |||||||||:::||||||||||||||||||    :::
Db   3101 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 3160

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
            |||   :::   ||||||      |||||||||   :::|||||||||      |||:::
Db   3161 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 3220

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
            ||||||||||||||||||||||||||||||      |||||||||:::      ||||||
Db   3221 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 3280

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
            :::||||||||||:::::::||||||   :::|||:::|||:::|||    :::|||||||||
Db   3281 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 3340

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
            ||||||||||||||||:::   |||:::||||||:::||||||||||||||||:::|||
Db   3341 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 3400

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
            ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3401 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 3460

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
            ||||||:::|||||||||||||||||||||||||||||:::|||||||:::::::
Db   3461 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 3520

Qy    501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
```

```
Db    3521 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 3580

Qy     521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db    3581 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 3640

Qy     541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db    3641 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 3700

Qy     561 Gln------------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
Db    3701 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 3760

Qy     575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 589
Db    3761 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 3820

Qy     590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
Db    3821 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 3880

Qy     608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627
Db    3881 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 3940

Qy     628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
Db    3941 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 4000

Qy     648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
Db    4001 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 4060

Qy     667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
Db    4061 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 4120

Qy     687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
Db    4121 ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 4180

Qy     707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
Db    4181 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 4240

Qy     727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
Db    4241 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 4300

Qy     747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
Db    4301 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 4360

Qy     767 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
Db    4361 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 4420

Qy     787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
Db    4421 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 4480

Qy     807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
Db    4481 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 4540

Qy     827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
Db    4541 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 4600

Qy     847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
Db    4601 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 4660

Qy     867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
Db    4661 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 4720

Qy     887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
Db    4721 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 4780

Qy     907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
Db    4781 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 4840

Qy     927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
Db    4841 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 4900

Qy     947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
Db    4901 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 4960

Qy     967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
Db    4961 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 5020

Qy     987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
Db    5021 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 5080

Qy    1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
Db    5081 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 5140

Qy    1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
Db    5141 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 5200

Qy    1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
Db    5201 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 5260

Qy    1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
Db    5261 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 5320

Qy    1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
Db    5321 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 5380

Qy    1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
Db    5381 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 5440

Qy    1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
Db    5441 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 5500

Qy    1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
Db    5501 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 5560

Qy    1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
Db    5561 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 5620

Qy    1187 LysLeuValAsnAlaGlnSer 1193
Db    5621 TCAATGGTCAGTGTCCAGGCT 5641
```

RESULT 3
US-08-793-610-6
; Sequence 6, Application US/08793610

```
; Patent No. 5858744
;  GENERAL INFORMATION:
;    APPLICANT:  BAUM, Christopher
;    APPLICANT:  STOCKING-HARBERS, Carol
;    APPLICANT:  OSTERTAG, Wolfram
;    TITLE OF INVENTION:  RETROVIRAL VECTOR HYBRIDS AND THE USE THEROF
;    TITLE OF INVENTION:  FOR GENE TRANSFER
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nikaido, Marmelstein, Murray & Oram LLP
;      STREET:  655 Fifteenth Street N.W. Suite 330
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/793,610
;      FILING DATE:  07-MAR-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P 44 31 973.8
;      FILING DATE:  08-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 195 03 952.1
;      FILING DATE:  07-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/03175
;      FILING DATE:  10-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berman, Richard J.
;      REGISTRATION NUMBER:  39,105
;      REFERENCE/DOCKET NUMBER:  P1614-7007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)638-5000
;      TELEFAX:  (202)638-4810
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9318 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA
US-08-793-610-6

Alignment Scores:
Pred. No.:               0           Length:         9318
Score:                   3545.00     Matches:        681
Percent Similarity:      75.55%      Conservative:   246
Best Local Similarity:   55.50%      Mismatches:     264
Query Match:             58.97%      Indels:         36
DB:                      2           Gaps:           9

US-09-873-409-6 (1-1195) x US-08-793-610-6 (1-9318)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    1926 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 1985

Qy      21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db    1986 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 2045

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db    2046 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 2105

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    2106 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 2165

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    2166 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 2225

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    2226 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 2285

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    2286 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 2345

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    2346 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 2405

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    2406 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 2465

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    2466 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 2525

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    2526 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 2585

Qy     201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db    2586 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 2645

Qy     221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db    2646 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 2705

Qy     241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db    2706 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 2759

Qy     261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db    2760 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 2819

Qy     281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db    2820 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 2879

Qy     301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db    2880 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 2939

Qy     321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db    2940 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 2999

Qy     341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db    3000 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 3059

Qy     361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db    3060 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 3119

Qy     381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db    3120 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 3179

Qy     401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db    3180 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 3239
```

```
Qy      421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
             :::||||||||::::::||||||     :::||||:::||:::||:    :::|||||||||
Db     3240 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 3299

Qy      441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
            |||||||||||||||||:::   |||:::||||||:::|||||||||||||||:::|||
Db     3300 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 3359

Qy      461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
            ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     3360 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 3419

Qy      481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
            ||||||||||||:::||||||||||||||||||||||||:::||||||::::::
Db     3420 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 3479

Qy      501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
            ||||||   ||||||:::||||||   |||||||||||||||||||||:::|||||||||
Db     3480 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 3539

Qy      521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
            |||||||||:::|||:::||||||:::|||          ||||||::::::    |||
Db     3540 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 3599

Qy      541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
            ||||||   |||   |||||||||   ::::::|||:::|||:::   ||||||
Db     3600 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 3659

Qy      561 Gln------------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
            |||                  :::::::::   |||   ::::::   :::|||
Db     3660 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 3719

Qy      575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 589
               :::   |||      :::|||          :::
Db     3720 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 3779

Qy      590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
               :::|||::::::   |||      |||:::      :::   ::::::
Db     3780 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 3839

Qy      608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627
            :::   |||:::|||   ||||||      :::|||:::|||||||||      ||||||
Db     3840 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 3899

Qy      628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
            |||:::   |||:::|||      :::::::::||||||   :::   |||   |||
Db     3900 CCTTATTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 3959

Qy      648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
            :::|||||||||:::|||||||||   :::|||         :::|||      |||
Db     3960 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 4019

Qy      667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
            :::   ::::::      ::::::|||::::::|||:::   |||||||||:::|||:::
Db     4020 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 4079

Qy      687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
            ::::::|||:::||||||      :::|||:::|||||||||||||||||||   ||||||
Db     4080 ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 4139

Qy      707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
            |||::::::   |||::::::||||||   ||||||::::::||||||:::   :::
Db     4140 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 4199

Qy      727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
            |||:::||||||   |||||||||   ||||||   |||   ||||||::::::||||||
Db     4200 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 4259

Qy      747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
               |||||||||:::   |||::::||||||||      |||:::|||      :::|||
Db     4260 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 4319

Qy      767 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
            |||||||||||||||||||::::::|||   |||:::|||:::|||   |||::::::
Db     4320 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 4379

Qy      787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
            |||:::   |||::::::|||      ::::::|||   |||   |||||||||:::
Db     4380 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 4439

Qy      807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
            ||||||:::   ||||||||||||||||||||||||:::||||||   |||||||::::||
Db     4440 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 4499

Qy      827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
            |||||||||:::|||:::   ||||||   ||||||   :::   ||||||      :::
Db     4500 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 4559

Qy      847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
            ||||||:::   :::||||||   |||   |||      ::::::|||:::   |||
Db     4560 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 4619

Qy      867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
            :::||||||:::|||         ||||||||||||||||||||||||:::
Db     4620 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 4679

Qy      887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
            |||:::   :::      :::|||:::|||   :::   :::||||||||||||:::
Db     4680 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 4739

Qy      907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
            |||:::               ||||||:::|||:::||||||      |||||||||:::
Db     4740 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 4799

Qy      927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
                  ::::::||||||   |||   |||||||||   |||   ||||||      |||
Db     4800 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 4859

Qy      947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
            :::|||   |||||||||:::   |||   ||||||   |||   ||||||   ||||||
Db     4860 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 4919

Qy      967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
            |||:::   |||:::|||::::||||||   |||   :::::::||||:::|||:::|||
Db     4920 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 4979

Qy      987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
            ||||||||||||||||||||||||||||||   ||||||||||||:::|||   ||||||
Db     4980 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 5039

Qy     1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
            |||:::   |||:::||||||   ||||||   :::   |||   |||||||||||||||
Db     5040 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 5099

Qy     1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
            ||||||:::   :::   ||||||   |||||||||:::||||||:::||||||||||||
Db     5100 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 5159

Qy     1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
            ||||||||||||||||||||||||||||||||||||         :::||||||   |||
Db     5160 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 5219

Qy     1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
            |||      ||||||||||||:::||||||||||   ||||||   |||||||||:::|||:::
Db     5220 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 5279

Qy     1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
            ||||||   ||||||   ||||||||||||||||||||||||:::||||||||||||
Db     5280 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 5339

Qy     1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
            |||||||:::::::::|||   |||||||||||||||||||||||||||||||||
Db     5340 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 5399

Qy     1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
            :::|||||||||||||||||||   |||||||||||||||||   ||||||||||||:::
```

```
Db    5400 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 5459

Qy    1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
Db    5460 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 5519

Qy    1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
Db    5520 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 5579

Qy    1187 LysLeuValAsnAlaGlnSer 1193
Db    5580 TCAATGGTCAGTGTCCAGGCT 5600
```

RESULT 4
US-08-784-649A-1
; Sequence 1, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
; APPLICANT: Sikic, Branimir I
; APPLICANT: Chen, Gang
; TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
; TITLE OF INVENTION: CYCLOSPORIN MODULATION
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 2200 Sand Hill Road
; CITY: Menlo Park
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/784,649A
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sherwood, Pamela J
; REGISTRATION NUMBER: Reg.No. 5830697 36,677
; REFERENCE/DOCKET NUMBER: 06037/007001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-322-5070
; TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4264 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-784-649A-1

Alignment Scores:

Pred. No.:	0	Length:	4264
Score:	3543.50	Matches:	682
Percent Similarity:	75.63%	Conservative:	246
Best Local Similarity:	55.58%	Mismatches:	262
Query Match:	58.94%	Indels:	37
DB:	2	Gaps:	10

US-09-873-409-6 (1-1195) x US-08-784-649A-1 (1-4264)

```
Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db     292 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 351

Qy      21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
Db     352 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 411

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db     412 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 471

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db     472 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 531

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db     532 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 591

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db     592 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 651

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db     652 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAAATT 711

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db     712 GGA---ATGTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 768

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db     769 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 828

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db     829 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 888

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db     889 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 948

Qy     201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db     949 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1008

Qy     221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db    1009 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1068

Qy     241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db    1069 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1122

Qy     261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db    1123 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1182

Qy     281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db    1183 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1242

Qy     301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db    1243 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1302

Qy     321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db    1303 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1362

Qy     341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db    1363 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1422

Qy     361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db    1423 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1482
```

```
Qy   381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db  1483 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1542

Qy   401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db  1543 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1602

Qy   421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db  1603 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1662

Qy   441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db  1663 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1722

Qy   461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db  1723 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1782

Qy   481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db  1783 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1842

Qy   501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db  1843 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 1902

Qy   521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db  1903 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 1962

Qy   541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db  1963 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 2022

Qy   561 Gln-----------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
Db  2023 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 2082

Qy   575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 589
Db  2083 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2142

Qy   590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
Db  2143 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2202

Qy   608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627
Db  2203 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2262

Qy   628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
Db  2263 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2322

Qy   648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
Db  2323 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2382

Qy   667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
Db  2383 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2442

Qy   687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
Db  2443 ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2502

Qy   707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
Db  2503 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2562

Qy   727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
Db  2563 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2622

Qy   747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
Db  2623 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2682

Qy   767 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
Db  2683 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 2742

Qy   787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
Db  2743 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 2802

Qy   807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
Db  2803 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 2862

Qy   827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
Db  2863 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2922

Qy   847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
Db  2923 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 2982

Qy   867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
Db  2983 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 3042

Qy   887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
Db  3043 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 3102

Qy   907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
Db  3103 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3162

Qy   927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
Db  3163 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3222

Qy   947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
Db  3223 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3282

Qy   967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
Db  3283 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3342

Qy   987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
Db  3343 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3402

Qy  1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
Db  3403 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3462

Qy  1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
Db  3463 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3522

Qy  1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
Db  3523 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATTGTGAGGGCA 3582

Qy  1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
Db  3583 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3642

Qy  1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
```

```
Db   3643 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3702

Qy   1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
          ||||||:::::::::|||   ||||||||||||||||||||||||||||||||||||
Db   3703 GCCCTTGTTAGACAGCCTCATATTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 3762

Qy   1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
          :::||||||||||||||||||   |||||||||||||||||   |||||||||||:::
Db   3763 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 3822

Qy   1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
          |||:::   |||||||||||   ||||||||||||||||||||||||||      |||
Db   3823 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 3882

Qy   1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
          |||::::::||||||   ||||||||||||:::||||||      :::   |||||||||
Db   3883 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 3942

Qy   1187 LysLeuValAsnAlaGlnSer 1193
             :::|||:::   |||:::
Db   3943 TCAATGGTCAGTGTCCAGGCT 3963

RESULT 5
US-08-784-649A-5
; Sequence 5, Application US/08784649A
; Patent No. 5830697
;  GENERAL INFORMATION:
;    APPLICANT:  Sikic, Branimir I
;    APPLICANT:  Chen, Gang
;    TITLE OF INVENTION:  P-GLYCOPROTEIN MUTANT RESISTANT TO
;    TITLE OF INVENTION:  CYCLOSPORIN MODULATION
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  2200 Sand Hill Road
;      CITY:  Menlo Park
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/784,649A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sherwood, Pamela J
;      REGISTRATION NUMBER:  Reg.No. 5830697 36,677
;      REFERENCE/DOCKET NUMBER:  06037/007001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-322-5070
;      TELEFAX:  415-854-0875
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4264 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-784-649A-5

Alignment Scores:
Pred. No.:              0               Length:         4264
Score:                  3543.50         Matches:        682
Percent Similarity:     75.63%          Conservative:   246
Best Local Similarity:  55.58%          Mismatches:     262
Query Match:            58.94%          Indels:         37
DB:                     2               Gaps:           10

US-09-873-409-6 (1-1195) x US-08-784-649A-5 (1-4264)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          |||::::::|||   ||||||::::::::::::||||||   |||||||||||   |||
Db    292 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 351

Qy     21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
          |||   |||||||||:::|||                           :::|||:::|||
Db    352 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 411

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                   ::::::      |||      ||||||
Db    412 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 471

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                |||         ||||||   |||||||||         |||:::      |||
Db    472 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 531

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
          ||||||:::|||   |||   :::   |||   ||||||      :::||||||||||||
Db    532 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 591

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
          |||||||||:::::::::   |||:::|||||||||||||||      |||:::||||||
Db    592 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 651

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
          ||||||||||||:::|||   |||:::   ||||||:::::::||||||||||||||||||
Db    652 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAAATT 711

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
             :::||||||:::|||:::||||||         |||         ||||||
Db    712 GGA---ATGTTCCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 768

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
          ||||||||||||||||||||   |||:::   ||||||::::::      |||||||||
Db    769 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 828

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
                ::::::::::::      |||   |||   |||||||||   ||||||:::||||||
Db    829 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 888

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
          ||||||||||||||||||||||||::::::||||||||||||||||||         |||
Db    889 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 948

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
          :::|||||||||:::||||||      :::||||||:::::::||||||      |||||||||
Db    949 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1008

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
          :::   |||   :::   :::|||:::|||||||||         :::         :::|||
Db   1009 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1068

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
             ||||||||||||||||||||||:::|||:::|||:::||||||      |||:::|||
Db   1069 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1122

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
          |||   ||||||   |||||||||||||||:::         ::::::   :::|||   |||
Db   1123 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1182

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
             |||      |||   ||||||   |||||||||||||||:::   ||||||::::::
Db   1183 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1242

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
          ||||||   |||||||||||||||:::::::||||   :::|||:::||||||:::::::|||
Db   1243 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1302

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
          :::|||   :::||||||:::||||||   |||:::||||||||||||      :::|||
Db   1303 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1362
```

```
Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   1363 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1422

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   1423 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1482

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   1483 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1542

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   1543 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1602

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1603 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1662

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1663 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1722

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   1723 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1782

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   1783 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1842

Qy    501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db   1843 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 1902

Qy    521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db   1903 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 1962

Qy    541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db   1963 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 2022

Qy    561 Gln------------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
Db   2023 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 2082

Qy    575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 589
Db   2083 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2142

Qy    590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
Db   2143 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2202

Qy    608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627
Db   2203 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2262

Qy    628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
Db   2263 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2322

Qy    648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
Db   2323 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2382

Qy    667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
Db   2383 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2442

Qy    687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
Db   2443 ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2502

Qy    707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
Db   2503 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2562

Qy    727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
Db   2563 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2622

Qy    747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
Db   2623 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2682

Qy    767 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
Db   2683 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 2742

Qy    787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
Db   2743 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 2802

Qy    807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
Db   2803 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 2862

Qy    827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
Db   2863 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 2922

Qy    847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
Db   2923 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 2982

Qy    867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
Db   2983 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 3042

Qy    887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
Db   3043 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 3102

Qy    907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
Db   3103 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3162

Qy    927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
Db   3163 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3222

Qy    947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
Db   3223 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3282

Qy    967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
Db   3283 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3342

Qy    987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
Db   3343 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3402

Qy   1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
Db   3403 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3462

Qy   1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
Db   3463 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3522

Qy   1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
```

```
Db     3523 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATTGTGAGGGCA 3582

Qy     1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
Db     3583 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3642

Qy     1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
Db     3643 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3702

Qy     1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
Db     3703 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 3762

Qy     1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
Db     3763 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 3822

Qy     1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
Db     3823 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 3882

Qy     1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
Db     3883 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 3942

Qy     1187 LysLeuValAsnAlaGlnSer 1193
Db     3943 TCAATGGTCAGTGTCCAGGCT 3963
```

RESULT 6
5206352-3
;Patent No. 5206352
; APPLICANT: Roninson, Igor B.;Pastan Ira H.;Gottesman,
;Michael M.
; TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA
;SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
; NUMBER OF SEQUENCES: 4
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/622,836
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 892,575
; FILING DATE: 01-AUG-1986
; APPLICATION NUMBER: 845,610
; FILING DATE: 28-MAR-1986
;SEQ ID NO:3:
; LENGTH: 4669
5206352-3

Alignment Scores:

Pred. No.:	0	Length:	4669
Score:	3543.00	Matches:	681
Percent Similarity:	75.47%	Conservative:	245
Best Local Similarity:	55.50%	Mismatches:	265
Query Match:	58.93%	Indels:	36
DB:	6	Gaps:	9

US-09-873-409-6 (1-1195) x 5206352-3 (1-4669)

```
Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db      575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy       21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db      635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy       32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db      695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy       44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db      755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy       62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db      815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy       82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db      875 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy      102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db      935 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATCAAGTTATTGGTGACAAAATT 994

Qy      121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db      995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy      141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db     1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy      161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db     1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy      181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db     1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy      201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db     1235 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1294

Qy      221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db     1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1354

Qy      241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db     1355 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1408

Qy      261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db     1409 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1468

Qy      281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db     1469 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528

Qy      301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db     1529 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1588

Qy      321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db     1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1648

Qy      341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db     1649 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1708

Qy      361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db     1709 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1768

Qy      381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db     1769 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1828

Qy      401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
```

```
Db   1829 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1888

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1889 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1948

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1949 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2008

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   2009 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2068

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   2069 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2128

Qy    501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db   2129 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 2188

Qy    521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db   2189 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 2248

Qy    541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db   2249 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 2308

Qy    561 Gln-----------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
Db   2309 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 2368

Qy    575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 589
Db   2369 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2428

Qy    590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
Db   2429 CGTAGGAGTGTCCGTGGATCACAAGCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2488

Qy    608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627
Db   2489 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2548

Qy    628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
Db   2549 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2608

Qy    648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
Db   2609 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2668

Qy    667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
Db   2669 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2728

Qy    687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
Db   2729 ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2788

Qy    707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
Db   2789 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2848

Qy    727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
Db   2849 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2908

Qy    747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
Db   2909 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2968

Qy    767 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
Db   2969 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 3028

Qy    787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
Db   3029 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 3088

Qy    807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
Db   3089 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 3148

Qy    827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
Db   3149 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 3208

Qy    847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
Db   3209 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 3268

Qy    867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
Db   3269 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 3328

Qy    887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
Db   3329 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 3388

Qy    907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
Db   3389 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3448

Qy    927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
Db   3449 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3508

Qy    947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
Db   3509 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3568

Qy    967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
Db   3569 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3628

Qy    987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
Db   3629 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3688

Qy   1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
Db   3689 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3748

Qy   1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
Db   3749 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3808

Qy   1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
Db   3809 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3868

Qy   1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
Db   3869 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3928

Qy   1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
Db   3929 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3988

Qy   1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
Db   3989 GCCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 4048
```

```
Qy    1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
Db    4049 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 4108

Qy    1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
Db    4109 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 4168

Qy    1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
Db    4169 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 4228

Qy    1187 LysLeuValAsnAlaGlnSer 1193
Db    4229 TCAATGGTCAGTGTCCAGGCT 4249
```

RESULT 7
US-08-752-447-1
; Sequence 1, Application US/08752447
; Patent No. 5994088
; GENERAL INFORMATION:
;   APPLICANT: Mechetner, Eugene
;   APPLICANT: Roninson, Igor B
;   TITLE OF INVENTION: Methods and Reagents for Preparing and
;   TITLE OF INVENTION: Using Immunoligcal Agents Specific for P-glycoprotein
;   NUMBER OF SEQUENCES: 2
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff Ltd.
;   STREET: 300 South Wacker Drive, Seventh Floor
;   CITY: Chicago
;   STATE: Illinois
;   COUNTRY: USA
;   ZIP: 60606
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/752,447
;   FILING DATE: 15-NOV-1996
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: No. 5994088nan, Kevin E
;   REGISTRATION NUMBER: 35,303
;   REFERENCE/DOCKET NUMBER: 95,1121
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 312-913-0001
;   TELEFAX: 312-913-9808
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4669 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: 1..424
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 425..4264
;   FEATURE:
;   NAME/KEY: 3'UTR
;   LOCATION: 4265..4669
US-08-752-447-1

Alignment Scores:

Pred. No.:	0	Length:	4669
Score:	3527.00	Matches:	677
Percent Similarity:	75.14%	Conservative:	245
Best Local Similarity:	55.18%	Mismatches:	269
Query Match:	58.67%	Indels:	36
DB:	2	Gaps:	9

US-09-873-409-6 (1-1195) x US-08-752-447-1 (1-4669)

```
Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db     575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy      21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db     635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe----------------- 43
Db     695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db     755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db     815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db     875 TTTTTTCATGCTATTATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db     935 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAAATT 994

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db     995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy     201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db    1235 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1294

Qy     221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db    1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1354

Qy     241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db    1355 GCTGTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1408

Qy     261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db    1409 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1468

Qy     281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db    1469 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528

Qy     301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db    1529 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1588

Qy     321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
```

```
Db   1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1648

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   1649 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1708

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   1709 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1768

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   1769 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1828

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   1829 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1888

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1889 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1948

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1949 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2008

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   2009 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2068

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   2069 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2128

Qy    501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db   2129 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 2188

Qy    521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db   2189 TTGCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 2248

Qy    541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db   2249 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 2308

Qy    561 Gln------------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
Db   2309 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 2368

Qy    575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 589
Db   2369 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2428

Qy    590 ---LysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLys 607
Db   2429 CGTAGGAGTGTCCGTGGATCACAAGCCCAACACAGAAAGCTTAGTACCAAAGAGGCTCTG 2488

Qy    608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627
Db   2489 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2548

Qy    628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
Db   2549 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2608

Qy    648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
Db   2609 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2668

Qy    667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
Db   2669 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2728

Qy    687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
Db   2729 ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2788

Qy    707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
Db   2789 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTCATGACCCTAAA 2848

Qy    727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
Db   2849 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2908

Qy    747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
Db   2909 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2968

Qy    767 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
Db   2969 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 3028

Qy    787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
Db   3029 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTTGCTGGACAAGCACTGAAAGATAAGAAA 3088

Qy    807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
Db   3089 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 3148

Qy    827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
Db   3149 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 3208

Qy    847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
Db   3209 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 3268

Qy    867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
Db   3269 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 3328

Qy    887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
Db   3329 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 3388

Qy    907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
Db   3389 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3448

Qy    927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
Db   3449 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3508

Qy    947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
Db   3509 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3568

Qy    967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
Db   3569 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3628

Qy    987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
Db   3629 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3688

Qy   1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
Db   3689 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3748

Qy   1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
Db   3749 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3808
```

```
Qy   1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
          ||||||||||||||||||||||||||||||||||||||        :::||||||     |||
Db   3809 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3868

Qy   1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
          |||      ||||||||||:::||||||||   ||||||   |||||:::|||:::
Db   3869 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3928

Qy   1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
          ||||||   ||||||   |||||||||||||||||||||||||||:::|||||||||||||
Db   3929 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3988

Qy   1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
             |||:::::::|||   ||||||||||||||||||||||||||||||||||||
Db   3989 CGCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 4048

Qy   1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
          :::||||||||||||||||||   ||||||||||||||||||   ||||||||||||:::
Db   4049 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 4108

Qy   1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
          |||:::   ||||||||||||   ||||||||||||||||||||||||||       |||
Db   4109 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 4168

Qy   1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
          |||::::::||||||   ||||||||||||:::||||||       :::   |||||||||
Db   4169 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 4228

Qy   1187 LysLeuValAsnAlaGlnSer 1193
             :::|||:::   |||:::
Db   4229 TCAATGGTCAGTGTCCAGGCT 4249
```

```
RESULT 8
US-09-316-167-1
; Sequence 1, Application US/09316167
; Patent No. 6365357
;  GENERAL INFORMATION:
;    APPLICANT:  Mechetner, Eugene
;    APPLICANT:  Roninson, Igor B
;    TITLE OF INVENTION:  Methods and Reagents for Preparing and
;    TITLE OF INVENTION:  Using Immunoligcal Agents Specific for P-glycoprotein
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McDonnell Boehnen Hulbert & Berghoff Ltd.
;      STREET:  300 South Wacker Drive, Seventh Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/316,167
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/752,447
;      FILING DATE:  15-NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6365357nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  95,1121
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-913-0001
;      TELEFAX:  312-913-9808
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4669 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  5'UTR
;      LOCATION:  1..424
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  425..4264
;    FEATURE:
;      NAME/KEY:  3'UTR
;      LOCATION:  4265..4669
US-09-316-167-1

Alignment Scores:
Pred. No.:               0              Length:         4669
Score:                   3527.00        Matches:        677
Percent Similarity:      75.14%         Conservative:   245
Best Local Similarity:   55.18%         Mismatches:     269
Query Match:             58.67%         Indels:         36
DB:                      4              Gaps:           9

US-09-873-409-6 (1-1195) x US-09-316-167-1 (1-4669)
```

```
Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          |||::::::|||   ||||||::::::::::::||||||   ||||||||||||   |||
Db    575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy     21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
          |||   |||||||||:::|||                           :::|||:::|||
Db    635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                   ::::::      |||      ||||||
Db    695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                |||         ||||||   |||||||||         |||:::      |||
Db    755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
          ||||||:::|||   |||   :::   |||   ||||||      :::||||||||||||
Db    815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
          ||||||||:::::::::   |||:::||||||||||||||||      |||:::||||||
Db    875 TTTTTTCATGCTATTATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
          ||||||||||||:::|||   |||:::   ||||||::::::||||||||||||||||||
Db    935 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAAATT 994

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
             :::   ||||||:::|||:::||||||       |||      ||||||      :::
Db    995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
          ||||||||||||||||||||||   |||:::   ||||||::::::      |||||||||
Db   1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
                ::::::::::::   |||   |||   |||||||||   ||||||:::||||||
Db   1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
          ||||||||||||||||||||||||::::::|||||||||||||||||||||      |||
Db   1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
          :::|||||||||:::||||||      :::||||||::::::||||||      |||||||||
Db   1235 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1294

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
```

```
Db   1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1354

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db   1355 GCTGTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1408

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db   1409 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1468

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db   1469 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db   1529 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1588

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db   1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1648

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   1649 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1708

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   1709 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1768

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   1769 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1828

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   1829 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1888

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1889 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1948

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1949 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2008

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   2009 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2068

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   2069 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2128

Qy    501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db   2129 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 2188

Qy    521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db   2189 TTGCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 2248

Qy    541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db   2249 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 2308

Qy    561 Gln------------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
Db   2309 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 2368

Qy    575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 589
Db   2369 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCCAGTCTAATAAGAAAAAGATCAACT 2428

Qy    590 ---LysSerIleLysSerAspPheIleAspLysAlaGluGluSerThr---GlnSerLys 607
Db   2429 CGTAGGAGTGTCCGTGGATCACAAGCCCAACACAGAAAGCTTAGTACCAAAGAGGCTCTG 2488

Qy    608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627
Db   2489 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2548

Qy    628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
Db   2549 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2608

Qy    648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
Db   2609 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2668

Qy    667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
Db   2669 CGACAGAATAGTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2728

Qy    687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
Db   2729 ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2788

Qy    707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
Db   2789 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTCATGACCCTAAA 2848

Qy    727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
Db   2849 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2908

Qy    747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
Db   2909 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAAATCTTGGGACAGGAATAATT 2968

Qy    767 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
Db   2969 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 3028

Qy    787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
Db   3029 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTTGCTGGACAAGCACTGAAAGATAAGAAA 3088

Qy    807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
Db   3089 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 3148

Qy    827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
Db   3149 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 3208

Qy    847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
Db   3209 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 3268

Qy    867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
Db   3269 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTC 3328

Qy    887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
Db   3329 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 3388

Qy    907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
Db   3389 GGGCAAGTCAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3448

Qy    927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
Db   3449 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3508
```

```
Qy     947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
               :::|||  |||||||||:::   |||  ||||||  |||  ||||||  ||||||
Db    3509 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3568

Qy     967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
           |||:::   :::|||:::||||||||||||   :::::::::|||:::|||:::|||
Db    3569 GACATCCCAGTGCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3628

Qy     987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
           |||||||||||||||||||||||||||||||   |||||||||||||:::|||  ||||||
Db    3629 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3688

Qy    1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
           |||:::   |||:::||||||  ||||||   :::  |||   |||||||||||||||||
Db    3689 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3748

Qy    1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
           ||||||:::   :::   ||||||  |||||||||:::||||||:::||||||||||||
Db    3749 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3808

Qy    1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
           ||||||||||||||||||||||||||||||||||||        :::||||||      |||
Db    3809 GAGAACATTGCCTATGGAGACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3868

Qy    1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
           |||     ||||||||||||:::|||||||||   ||||||   ||||||:::|||:::
Db    3869 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3928

Qy    1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
           ||||||   ||||||   ||||||||||||||||||||||||||||:::||||||||||||
Db    3929 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3988

Qy    1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
                  |||:::::::|||   |||||||||||||||||||||||||||||||||
Db    3989 CGCCTTGTTAGACAGCCTCATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 4048

Qy    1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
           :::||||||||||||||||||   ||||||||||||||||||   ||||||||||||:::
Db    4049 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAAGCCAGAGAAGGCCGCACCTGCATT 4108

Qy    1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
           |||:::   ||||||||||||   ||||||||||||||||||||||||||||      |||
Db    4109 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 4168

Qy    1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
           |||::::::||||||   ||||||||||||:::||||||      :::   |||||||||
Db    4169 GGCAGAGTCAAGGAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 4228

Qy    1187 LysLeuValAsnAlaGlnSer 1193
              :::|||:::   |||:::
Db    4229 TCAATGGTCAGTGTCCAGGCT 4249
```

RESULT 9
US-09-120-513-1
; Sequence 1, Application US/09120513
; Patent No. 6025160
; GENERAL INFORMATION:
;   APPLICANT: Brun, Kimberly
;   APPLICANT: Chenery, Richard
;   APPLICANT: Ellens, Harma
;   APPLICANT: Field, John
;   APPLICANT: Yue, Lin
;   TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE
;   TITLE OF INVENTION: SEQUENCES ENCODING RAT MDR1B2 AND
;   TITLE OF INVENTION: SCREENING METHODS THEREOF
;   NUMBER OF SEQUENCES: 2
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: SmithKline Beecham Corporation
;   STREET: 709 Swedeland Road
;   CITY: King of Prussia
;   STATE: PA
;   COUNTRY:
;   ZIP: 19406
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Diskette
;   COMPUTER: IBM Compatible
;   OPERATING SYSTEM: DOS
;   SOFTWARE: FastSEQ for Windows Version 2.0
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/120,513
;   FILING DATE: 22-JUL-1998
;   CLASSIFICATION:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: King, William T
;   REGISTRATION NUMBER: 30,954
;   REFERENCE/DOCKET NUMBER: GP50008
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 610-270-5015
;   TELEFAX: 610-270-5090
;   TELEX:
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4233 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: Genomic DNA
US-09-120-513-1

```
Alignment Scores:
Pred. No.:              0              Length:         4233
Score:                  3517.50        Matches:        684
Percent Similarity:     74.47%         Conservative:   229
Best Local Similarity:  55.79%         Mismatches:     274
Query Match:            58.51%         Indels:         39
DB:                     3              Gaps:           9
```

US-09-873-409-6 (1-1195) x US-09-120-513-1 (1-4233)

```
Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||  ||||||   ||||||:::::::::::::|||        |||||||||:::   |||
Db     170 ATGGCTCTGGGAACTCTCGCTGCTATCATCCACGGAACCCTGCTTCCCCTCCTGATGCTG 229

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
           |||   |||   |||:::|||:::
Db     230 GTGTTCGGATACATGACAGATAGTTTTACCCAAGCAGAGACCCGCATTCTGCCGAGCGTT 289

Qy      31 ------SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu--------- 45
                 :::       |||   |||   |||            |||
Db     290 ACTAATCAAAGTGAAATCAACAGTACACAGACCGTCAGCGACAGCAGTCTGGAGGAGGAC 349

Qy      46 -----------ThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                      ||||||   |||||||||            ||||||        |||
Db     350 ATGGCCATGTACGCCTACTATTACACGGGCATTGGTGCCGGTGTGCTCATCGTTGCCTAC 409

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
           ||||||:::|||||||||   :::   |||   ||||||        :::|||||||::::::
Db     410 ATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAAG 469

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
           |||||||||:::::::::   |||:::|||||||||||||||      |||   ||||||
Db     470 TTTTTCCATGCCATCATGAATCAGGAGATAGGCTGGTTTGACGTGAATGACGCTGGGGAG 529

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
           ||||||||||||:::|||   |||:::   ||||||:::||||||||||||||||||:::
Db     530 CTCAACACCCGGCTCACAGATGACGTCTCCAAAATTAATGACGGAATTGGTGACAAACTT 589

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
              :::   ||||||:::::::::|||||||||   |||      :::|||   :::
Db     590 GGAATGTTCTTTCAGTCCATAACGACATTTTCAGCCGGTTTTATAATAGGATTTATAAGT 649

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
           |||||||||||||||||||   |||:::   ||||||||||||      |||:::|||
Db     650 GGTTGGAAGCTAACCCTTGTAATTTGGCCGTCAGCCCTCTTATTGGGTTGTCATCTGCC 709
```

```
Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    710 ATGTGGGCAAAGGTACTGACTTCATTTACTAATAAGGAACTCCAGGCTTATGCGAAAGCT 769
Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    770 GGAGCAGTTGCCGAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCGTTTGGAGGACAA 829
Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db    830 AAGAAGGAACTTGAAAGGTACAATAAAAATTTAGAAGAAGCTAAAAGAGTTGGCATAAAG 889
Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db    890 AAAGCCATCACGGCCAACATTTCCATAGGTATTGCCTACCTGTTGGTCTATGCGTCTTAT 949
Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db    950 GCACTGGCATTCTGGTATGGGACCTCCTTGGTCCTCTCAAATGAA------TATTCTATT 1003
Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db   1004 GGACAAGTGCTTACCGTCTTCTTCTCTATTTTATTGGGGACTTTCAGTATTGGACATTTA 1063
Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db   1064 GCCCCAAACATAGAAGCCTTTGCAAATGCAAGAGGGGCAGCCTATGAAATCTTCAAGATA 1123
Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db   1124 ATTGATAATGAGCCAAGCATTGACAGCTTCTCAACCAAGGGACACAAACCAGACAGTATA 1183
Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db   1184 ATGGGAAATTTGGAATTTAAAAATGTTTACTTCAACTACCCATCACGAAGTGAAGTTAAG 1243
Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   1244 ATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTGGTTGGCAAC 1303
Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   1304 AGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCATAGAG 1363
Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   1364 GGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAA 1423
Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   1424 ATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCCACCACGATTGCCGAAAACATT 1483
Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1484 CGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAAT 1543
Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1544 GCCTATGACTTCATCATGAAACTGCCCCACAAATTTAACACCCTGGTTGGTGAGAGAGGG 1603
Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   1604 GCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAAC 1663
Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   1664 CCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTG 1723
Qy    501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db   1724 GTTCAGGCCGCTCTGGATAAGGCTAGAGAAGGCCGGACCACCATTGTGATAGCTCACCGC 1783
Qy    521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db   1784 TTGTCTACAGTGCGCAATGCTGACGTCATTGCTGGTTTTGATGGTGGTGTCATTGTGGAG 1843
Qy    541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db   1844 CAAGGAAATCATGAAGAGCTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTCATGACA 1903
Qy    561 GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 580
Db   1904 CAG---ACTAGAGGAAATGAAATTGAACCAGGAAATAATGCTTATGAATCCCAAAGTGAC 1960
Qy    581 ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAla 600
Db   1961 ACTGGTGCCTCTGAGTTGACTTCAGAAAAATCA---AAATCTCCTTTAATAAGGAGATCA 2017
Qy    601 ---------------------------------GluGluSerThrGlnSerLysGlu--- 608
Db   2018 ATTCGCAGAAGTATCCACAGAAGACAAGACCAGGAGAGAAGACTTAGTTCGAAAGAGGAT 2077
Qy    609 ------IleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGlu 626
Db   2078 GTGGATGAAGATGTGCCTATGGTTTCCTTTTGGCAGATCCTAAAGCTAAATATTAGTGAA 2137
Qy    627 TrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProVal 646
Db   2138 TGGCCCTATTTAGTTGTGGGTGTACTTTGTGCTGTTATAAATGGGTGCATACAACCAGTG 2197
Qy    647 PheSerIleIlePheAlaLysIleIleThrMetPheGly---AsnAsnAspLysThrThr 665
Db   2198 TTTGCCATAGTGTTTTCAAAGATTGTAGGGGTTTTTTCAAGAGACGACGACCATGAAACC 2257
Qy    666 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 685
Db   2258 AAACAACGGAATTGTAACTTGTTTTCCCTTCTCTTTCTGGTCATGGGAATGATTTCTTTT 2317
Qy    686 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 705
Db   2318 GTTACGTACTTCTTTCAAGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGA 2377
Qy    706 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLys 725
Db   2378 CTCCGATACATGGTCTTCAAATCCATGCTGCGACAGGATATAAGCTGGTTTGATGACCAT 2437
Qy    726 GluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 745
Db   2438 AAAAACACCACTGGCTCGCTGACTACCAGGCTCGCTAGTGACGCTTCTAATGTTAAAGGG 2497
Qy    746 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 765
Db   2498 GCTATGGGCTCCAGGCTTGCTGTAGTTACCCAGAATGTAGCAAACCTTGGCACAGGAATT 2557
Qy    766 IleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVal 785
Db   2558 ATCTTATCCTTAGTCTATGGCTGGCAGCTTACACTTTTACTTGTAGTAATTATACCACTC 2617
Qy    786 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 805
Db   2618 ATTGTCTTGGGTGGAATTATTGAAATGAAACTGTTGTCTGGTCAAGCCTTGAAGGACAAG 2677
Qy    806 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIle 825
Db   2678 AAAGAGCTAGAGATCTCTGGGAAGATCGCTACAGAAGCAATTGAAAACTTCCGCACTGTT 2737
Qy    826 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGln 845
Db   2738 GTCTCTTTGACTAGGGAGCAGAAGTTTGAAACTATGTATGCCCAGAGCTTGCAGATACCA 2797
Qy    846 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla 865
Db   2798 TACAGAAATGCTTTGAAGAAAGCACACGTCTTTGGGATCACCTTCGCCTTCACCCAGGCC 2857
Qy    866 PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGly 885
```

```
Db   2858 ATGATTTATTTTTCCTATGCTGCTTGTTTCCGGTTCGGTGCCTACTTGGTGGCACGAGAA 2917

Qy    886 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAla 905
Db   2918 CTCATGACGTTTGAAAATGTTATGTTGGTATTTTCTGCTGTTGTCTTTGGTGCCATGGCA 2977

Qy    906 IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHis 925
Db   2978 GCAGGGAATACCAGTTCATTCGCTCCTGACTACGCGAAGGCCAAAGTCTCAGCATCCCAC 3037

Qy    926 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLys 945
Db   3038 ATCATCAGGATCATTGAGAAAATCCCCGAGATTGACAGCTACAGCACGGAGGGCTTGAAG 3097

Qy    946 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 965
Db   3098 CCTAATTGGTTAGAAGGAAATGTGAAATTTAATGGAGTCATGTTCAACTATCCCACCCGA 3157

Qy    966 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 985
Db   3158 CCCAACATCCCAGTGCTTCAGGGACTGAGCTTCGAGGTGAAGAAGGGGCAAACGCTTCGC 3217

Qy    986 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr 1005
Db   3218 CTGGTGGGCAGCAGTGGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTCGAGCGCTTCTAC 3277

Qy   1006 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 1025
Db   3278 AACCCCATGGCTGGAACAGTGTTTCTAGATGGCAAAGAAATAAAACAACTCAACGTCCAG 3337

Qy   1026 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 1045
Db   3338 TGGCTCCGCGCCCACCTGGGCATTGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATC 3397

Qy   1046 AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu 1065
Db   3398 ACCGAGAACATCGCCTACGGAGACAACAGCCGTGTCGTGTCTCATGAGGAGATCGTGAGG 3457

Qy   1066 AlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThr 1085
Db   3458 GCCGCCAGGGAGGCCAACATCCACCAGTTCATCGACTCACTGCCTGAGAAATACAACACC 3517

Qy   1086 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAla 1105
Db   3518 AGAGTGGGAGACAAAGGGACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATCGCCATCGCG 3577

Qy   1106 ArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsp 1125
Db   3578 CGCGCCCTCGTCAGACAGCCTCACATCTTACTTCTGGATGAAGCGACATCAGCTCTGGAT 3637

Qy   1126 AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys 1145
Db   3638 ACGGAGAGTGAAAAGGTCGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACCTGC 3697

Qy   1146 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 1165
Db   3698 ATTGTGATCGCGCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATTCAG 3757

Qy   1166 AsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyr 1185
Db   3758 AACGGCCAGGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCCCAGAAAGGCATCTAT 3817

Qy   1186 PheLysLeuValAsnAla 1191
Db   3818 TTCTCGATGGTTCAGGCT 3835
```

RESULT 10
US-09-450-105-1
; Sequence 1, Application US/09450105
; Patent No. 6169166
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: ENCODING RAT MDR1B2 AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP-50008-D1
; CURRENT APPLICATION NUMBER: US/09/450,105
; CURRENT FILING DATE: 1999-11-29
; EARLIER APPLICATION NUMBER: 09/120,513
; EARLIER FILING DATE: 1998-07-22
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4233
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
US-09-450-105-1

```
Alignment Scores:
Pred. No.:               0          Length:        4233
Score:                   3517.50    Matches:       684
Percent Similarity:      74.47%     Conservative:  229
Best Local Similarity:   55.79%     Mismatches:    274
Query Match:             58.51%     Indels:        39
DB:                      4          Gaps:          9
```

US-09-873-409-6 (1-1195) x US-09-450-105-1 (1-4233)

```
Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    170 ATGGCTCTGGGAACTCTCGCTGCTATCATCCACGGAACCCTGCTTCCCCTCCTGATGCTG 229

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
Db    230 GTGTTCGGATACATGACAGATAGTTTTACCCAAGCAGAGACCCGCATTCTGCCGAGCGTT 289

Qy     31 ------SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu--------- 45
Db    290 ACTAATCAAAGTGAAATCAACAGTACACAGACCGTCAGCGACAGCAGTCTGGAGGAGGAC 349

Qy     46 -----------ThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    350 ATGGCCATGTACGCCTACTATTACACGGGCATTGGTGCCGGTGTGCTCATCGTTGCCTAC 409

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    410 ATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAAG 469

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    470 TTTTTCCATGCCATCATGAATCAGGAGATAGGCTGGTTTGACGTGAATGACGCTGGGGAG 529

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    530 CTCAACACCCGGCTCACAGATGACGTCTCCAAAATTAATGACGGAATTGGTGACAAACTT 589

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    590 GGAATGTTCTTTCAGTCCATAACGACATTTTCAGCCGGTTTTATAATAGGATTTATAAGT 649

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    650 GGTTGGAAGCTAACCCTTGTAATTTTGGCCGTCAGCCCTCTTATTGGGTTGTCATCTGCC 709

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    710 ATGTGGGCAAAGGTACTGACTTCATTTACTAATAAGGAACTCCAGGCTTATGCGAAAGCT 769

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    770 GGAGCAGTTGCCGAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCGTTTGGAGGACAA 829

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
```

```
Db    830 AAGAAGGAACTTGAAAGGTACAATAAAAATTTAGAAGAAGCTAAAAGAGTTGGCATAAAG 889

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db    890 AAAGCCATCACGGCCAACATTTCCATAGGTATTGCCTACCTGTTGGTCTATGCGTCTTAT 949

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db    950 GCACTGGCATTCTGGTATGGGACCTCCTTGGTCCTCTCAAATGAA------TATTCTATT 1003

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db   1004 GGACAAGTGCTTACCGTCTTCTTCTCTATTTTATTGGGGACTTTCAGTATTGGACATTTA 1063

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db   1064 GCCCCAAACATAGAAGCCTTTGCAAATGCAAGAGGGGCAGCCTATGAAATCTTCAAGATA 1123

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db   1124 ATTGATAATGAGCCAAGCATTGACAGCTTCTCAACCAAGGGACACAAACCAGACAGTATA 1183

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db   1184 ATGGGAAATTTGGAATTTAAAAATGTTTACTTCAACTACCCATCACGAAGTGAAGTTAAG 1243

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   1244 ATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTGGTTGGCAAC 1303

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   1304 AGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCATAGAG 1363

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   1364 GGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAA 1423

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   1424 ATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCCACCACGATTGCCGAAAACATT 1483

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1484 CGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAAT 1543

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1544 GCCTATGACTTCATCATGAAACTGCCCCACAAATTTAACACCCTGGTTGGTGAGAGAGGG 1603

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   1604 GCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAAC 1663

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   1664 CCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTG 1723

Qy    501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db   1724 GTTCAGGCCGCTCTGGATAAGGCTAGAGAAGGCCGGACCACCATTGTGATAGCTCACCGC 1783

Qy    521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db   1784 TTGTCTACAGTGCGCAATGCTGACGTCATTGCTGGTTTTGATGGTGGTGTCATTGTGGAG 1843

Qy    541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer\560
Db   1844 CAAGGAAATCATGAAGAGCTCATGAAAGAGAAGGGCATTTACTTCAAACTTGTCATGACA 1903

Qy    561 GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 580
Db   1904 CAG---ACTAGAGGAAATGAAATTGAACCAGGAAATAATGCTTATGAATCCCAAAGTGAC 1960

Qy    581 ThrAsnSerLeuProLeuHisSerValLysSerIleLysSerAspPheIleAspLysAla 600
Db   1961 ACTGGTGCCTCTGAGTTGACTTCAGAAAAATCA---AAATCTCCTTTAATAAGGAGATCA 2017

Qy    601 --------------------------------GluGluSerThrGlnSerLysGlu--- 608
Db   2018 ATTCGCAGAAGTATCCACAGAAGACAAGACCAGGAGAGAAGACTTAGTTCGAAAGAGGAT 2077

Qy    609 ------IleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGlu 626
Db   2078 GTGGATGAAGATGTGCCTATGGTTTCCTTTTGGCAGATCCTAAAGCTAAATATTAGTGAA 2137

Qy    627 TrpProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProVal 646
Db   2138 TGGCCCTATTTAGTTGTGGGTGTACTTTGTGCTGTTATAAATGGGTGCATACAACCAGTG 2197

Qy    647 PheSerIleIlePheAlaLysIleIleThrMetPheGly---AsnAsnAspLysThrThr 665
Db   2198 TTTGCCATAGTGTTTTCAAAGATTGTAGGGGTTTTTTCAAGAGACGACGACCATGAAACC 2257

Qy    666 LeuLysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPhe 685
Db   2258 AAACAACGGAATTGTAACTTGTTTTCCCTTCTCTTTCTGGTCATGGGAATGATTTCTTTT 2317

Qy    686 ValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArg 705
Db   2318 GTTACGTACTTCTTTCAAGGCTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGA 2377

Qy    706 LeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLys 725
Db   2378 CTCCGATACATGGTCTTCAAATCCATGCTGCGACAGGATATAAGCTGGTTTGATGACCAT 2437

Qy    726 GluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGly 745
Db   2438 AAAAACACCACTGGCTCGCTGACTACCAGGCTCGCTAGTGACGCTTCTAATGTTAAAGGG 2497

Qy    746 AlaThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerVal 765
Db   2498 GCTATGGGCTCCAGGCTTGCTGTAGTTACCCAGAATGTAGCAAACCTTGGCACAGGAATT 2557

Qy    766 IleIleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProVal 785
Db   2558 ATCTTATCCTTAGTCTATGGCTGGCAGCTTACACTTTTACTTGTAGTAATTATACCACTC 2617

Qy    786 LeuAlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLys 805
Db   2618 ATTGTCTTGGGTGGAATTATTGAAATGAAACTGTTGTCTGGTCAAGCCTTGAAGGACAAG 2677

Qy    806 GlnGluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIle 825
Db   2678 AAAGAGCTAGAGATCTCTGGGAAGATCGCTACAGAAGCAATTGAAAACTTCCGCACTGTT 2737

Qy    826 ValSerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGln 845
Db   2738 GTCTCTTTGACTAGGGAGCAGAAGTTTGAAACTATGTATGCCCAGAGCTTGCAGATACCA 2797

Qy    846 HisArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAla 865
Db   2798 TACAGAAATGCTTTGAAGAAAGCACACGTCTTTGGGATCACCTTCGCCTTCACCCAGGCC 2857

Qy    866 PheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGly 885
Db   2858 ATGATTTATTTTTCCTATGCTGCTTGTTTCCGGTTCGGTGCCTACTTGGTGGCACGAGAA 2917

Qy    886 ArgMetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAla 905
Db   2918 CTCATGACGTTTGAAAATGTTATGTTGGTATTTTCTGCTGTTGTCTTTGGTGCCATGGCA 2977

Qy    906 IleGlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHis 925
Db   2978 GCAGGGAATACCAGTTCATTCGCTCCTGACTACGCGAAGGCCAAAGTCTCAGCATCCCAC 3037
```

```
Qy     926 LeuPheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLys 945
           :::       :::::||||||   |||   ||||||||   |||   ||||||    |||
Db    3038 ATCATCAGGATCATTGAGAAAATCCCCGAGATTGACAGCTACAGCACGGAGGGCTTGAAG 3097

Qy     946 ProAspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArg 965
           |||:::       |||||||||::::::|||      |||   |||   ||||||   |||
Db    3098 CCTAATTGGTTAGAAGGAAATGTGAAATTTAATGGAGTCATGTTCAACTATCCCACCCGA 3157

Qy     966 ProAspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAla 985
           |||::::::   :::|||:::||||||||         :::::::::|||:::|||:::
Db    3158 CCCAACATCCCAGTGCTTCAGGGACTGAGCTTCGAGGTGAAGAAGGGGCAAACGCTTCGC 3217

Qy     986 PheValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyr 1005
              |||||||||||||||||||||||||||||||||   |||||||||||:::|||   |||
Db    3218 CTGGTGGGCAGCAGTGGCTGCGGGAAGAGTACAGTGGTCCAGCTGCTCGAGCGCTTCTAC 3277

Qy    1006 AspProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGln 1025
           :::|||:::      |||   |||         ||||||   :::   |||:::||||||||||||
Db    3278 AACCCCATGGCTGGAACAGTGTTTCTAGATGGCAAAGAAATAAAACAACTCAACGTCCAG 3337

Qy    1026 TrpLeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIle 1045
           |||||||||:::   :::   ||||||   |||||||||:::||||||:::|||||||||
Db    3338 TGGCTCCGCGCCCACCTGGGCATTGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATC 3397

Qy    1046 AlaGluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGlu 1065
              |||||||||||||||||||||||||||||||||||||||         :::||||||
Db    3398 ACCGAGAACATCGCCTACGGAGACAACAGCCGTGTCGTGTCTCATGAGGAGATCGTGAGG 3457

Qy    1066 AlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThr 1085
           ||||||      ||||||||||||   ||||||:::   ||||||||||||||||||||||||
Db    3458 GCCGCCAGGGAGGCCAACATCCACCAGTTCATCGACTCACTGCCTGAGAAATACAACACC 3517

Qy    1086 GlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAla 1105
           :::||||||      ||||||   |||||||||||||||||||||||||:::|||||||||
Db    3518 AGAGTGGGAGACAAAGGGACTCAGCTGTCGGGCGGGCAGAAGCAGCGCATCGCCATCGCG 3577

Qy    1106 ArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAsp 1125
           |||||||||:::::::::|||   |||||||||||||||||||||||||||||||||||||
Db    3578 CGCGCCCTCGTCAGACAGCCTCACATCTTACTTCTGGATGAAGCGACATCAGCTCTGGAT 3637

Qy    1126 AsnAspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCys 1145
              :::||||||||||||||||||   ||||||||||||||||||   ||||||||||||
Db    3638 ACGGAGAGTGAAAAGGTCGTCCAGGAAGCGCTGGACAAAGCCAGGGAAGGCCGCACCTGC 3697

Qy    1146 LeuValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHis 1165
           :::|||:::   ||||||||||||   ||||||||||||||||||||||||||:::
Db    3698 ATTGTGATCGCGCACCGCCTGTCCACCATCCAGAACGCAGACTTGATCGTGGTGATTCAG 3757

Qy    1166 AsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyr 1185
           ||||||::::::||||||   ||||||||||||:::||||||         :::   ||||||
Db    3758 AACGGCCAGGTCAAGGAGCACGGCACCCACCAGCAGCTGCTGGCCCAGAAAGGCATCTAT 3817

Qy    1186 PheLysLeuValAsnAla 1191
           |||   :::|||   |||
Db    3818 TTCTCGATGGTTCAGGCT 3835
```

RESULT 11
US-08-583-276-18
; Sequence 18, Application US/08583276
; Patent No. 5837536
; GENERAL INFORMATION:
;   APPLICANT: McDonagh, Kevin T.
;   APPLICANT: Nienhuis, Arthur
;   APPLICANT: Tolstoshev, Paul
;   TITLE OF INVENTION: IMPROVED EXPRESSION OF HUMAN
;   TITLE OF INVENTION: MULTIDRUG RESISTANCE GENES AND IMPROVED
;   TITLE OF INVENTION: SELECTION OF CELLS TRANSDUCED WITH SUCH GENES
;   NUMBER OF SEQUENCES: 19
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Carella, Byrne, Bain, Gilfillan,
;   ADDRESSEE: Cecchi & Stewart
;   STREET: 6 Becker Farm Road
;   CITY: Roseland
;   STATE: New Jersey
;   COUNTRY: USA
;   ZIP: 07068
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: 3.5 inch diskette
;   COMPUTER: IBM PS/2
;   OPERATING SYSTEM: PC-DOS
;   SOFTWARE: DW4.V2
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/583,276
;   FILING DATE: 05-JAN-1996
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/332,444
;   FILING DATE: 31-OCT-1994
;   APPLICATION NUMBER: 07/887,712
;   FILING DATE: 22-MAY-1992
; INFORMATION FOR SEQ ID NO: 18:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4669 bases
;   TYPE: nucleic acid
;   STRANDEDNESS: singular
;   TOPOLOGY: linear
;   MOLECULE TYPE:
;   DESCRIPTION: Genomic DNA
US-08-583-276-18

Alignment Scores:

Pred. No.:	0	Length:	4669
Score:	3497.00	Matches:	673
Percent Similarity:	74.96%	Conservative:	246
Best Local Similarity:	54.89%	Mismatches:	271
Query Match:	58.17%	Indels:	36
DB:	2	Gaps:	9

US-09-873-409-6 (1-1195) x US-08-583-276-18 (1-4669)

```
Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||::::::|||   ||||||::::::::::::||||||   ||||||||||||   |||
Db     575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy      21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
           |||   |||||||||:::|||                          :::|||:::|||
Db     635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                    ::::::         |||      ||||||
Db     695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                 |||         ||||||   |||||||||         |||:::      |||
Db     755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
           ||||||:::|||   |||   :::   |||   ||||||         :::||||||||||||
Db     815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
           |||||||||:::::::::      |||:::||||||||||||||||||      |||:::|||:::
Db     875 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGCAG 934

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
           ||||||||||||:::|||      |||:::      ||||||::::::   |||||||||||||||
Db     935 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAACTTATTGGTGACAAAATT 994

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
              :::   ||||||:::|||:::||||||         |||      ||||||      :::
Db     995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
           |||||||||||||||||||||   |||:::   ||||||::::::      |||||||||
Db    1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114
```

```
Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db   1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db   1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db   1235 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1294

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db   1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1354

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db   1355 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1408

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db   1409 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1468

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db   1469 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db   1529 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1588

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db   1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1648

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   1649 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTCCCCCTGGTTGGAAAC 1708

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   1709 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1768

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   1769 GGGATGGTCAGTGTTCATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1828

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   1829 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1888

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1889 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1948

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1949 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2008

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   2009 CCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2068

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   2069 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2128

Qy    501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db   2129 GTTCAGGTGGCTCTGGATAAGGCCAGAAAAGGTCGGACCACCATTGTGATAGCTCATCGT 2188

Qy    521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db   2189 TTGTCTACAGTTCGTAATGCTGACGTCATCGCTGGTTTCGATGATGGAGTCATTGTGGAG 2248

Qy    541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db   2249 AAAGGAAATCATGATGAACTCATGAAAGAGAAAGGCATTTACTTCAAACTTGTCACAATG 2308

Qy    561 Gln-----------------AspIleLysLysAlaAspGluGlnMetGluSerMetThr 574
Db   2309 CAGACAGCAGGAAATGAAGTTGAATTAGAAAATGCAGCTGATGAATCCAAAAGTGAAATT 2368

Qy    575 TyrSerThrGluArgLysThrAsnSerLeuProLeuHisSerVal--------------- 589
Db   2369 GATGCCTTGGAAATGTCTTCAAATGATTCAAGATCACGTCTAATAAGAAAAAGATCAACT 2428

Qy    590 ---LysSerIleLys---SerAspPheIleAspLysAlaGluGluSerThrGlnSerLys 607
Db   2429 CGTAGGAGTGTCCGTGGATCACAACCCCAAGACAGAAAGCTTAGTACCAAAGAGGCTCTG 2488

Qy    608 GluIleSerLeuProGluValSerLeuLeuLysIleLeuLysLeuAsnLysProGluTrp 627
Db   2489 GATGAAAGTATACCTCCAGTTTCCTTTTGGAGGATTATGAAGCTAAATTTAACTGAATGG 2548

Qy    628 ProPheValValLeuGlyThrLeuAlaSerValLeuAsnGlyThrValHisProValPhe 647
Db   2549 CCTTATTTTGTTGTTGGTGTATTTTGTGCCATTATAAATGGAGGCCTGCAACCAGCATTT 2608

Qy    648 SerIleIlePheAlaLysIleIleThrMetPheGlyAsn---AsnAspLysThrThrLeu 666
Db   2609 GCAATAATATTTTCAAAGATTATAGGGGTTTTTACAAGAATTGATGATCCTGAAACAAAA 2668

Qy    667 LysHisAspAlaGluIleTyrSerMetIlePheValIleLeuGlyValIleCysPheVal 686
Db   2669 CGACAGAATACTAACTTGTTTTCACTATTGTTTCTAGCCCTTGGAATTATTTCTTTTATT 2728

Qy    687 SerTyrPheMetGlnGlyLeuPheTyrGlyArgAlaGlyGluIleLeuThrMetArgLeu 706
Db   2729 ACATTTTTCCTTCAGGGTTTCACATTTGGCAAAGCTGGAGAGATCCTCACCAAGCGGCTC 2788

Qy    707 ArgHisLeuAlaPheLysAlaMetLeuTyrGlnAspIleAlaTrpPheAspGluLysGlu 726
Db   2789 CGATACATGGTTTTCCGATCCATGCTCAGACAGGATGTGAGTTGGTTTGATGACCCTAAA 2848

Qy    727 AsnSerThrGlyGlyLeuThrThrIleLeuAlaIleAspIleAlaGlnIleGlnGlyAla 746
Db   2849 AACACCACTGGAGCATTGACTACCAGGCTCGCCAATGATGCTGCTCAAGTTAAAGGGGCT 2908

Qy    747 ThrGlySerArgIleGlyValLeuThrGlnAsnAlaThrAsnMetGlyLeuSerValIle 766
Db   2909 ATAGGTTCCAGGCTTGCTGTAATTACCCAGAATATAGCAATTCTTGGGACAGGAATAATT 2968

Qy    767 IleSerPheIleTyrGlyTrpGluMetThrPheLeuIleLeuSerIleAlaProValLeu 786
Db   2969 ATATCCTTCATCTATGGTTGGCAACTAACACTGTTACTCTTAGCAATTGTACCCATCATT 3028

Qy    787 AlaValThrGlyMetIleGluThrAlaAlaMetThrGlyPheAlaAsnLysAspLysGln 806
Db   3029 GCAATAGCAGGAGTTGTTGAAATGAAAATGTTGTCTGGACAAGCACTGAAAGATAAGAAA 3088

Qy    807 GluLeuLysHisAlaGlyLysIleAlaThrGluAlaLeuGluAsnIleArgThrIleVal 826
Db   3089 GAACTAGAAGGTGCTGGGAAGATCGCTACTGAAGCAATAGAAAACTTCCGAACCGTTGTT 3148

Qy    827 SerLeuThrArgGluLysAlaPheGluGlnMetTyrGluGluMetLeuGlnThrGlnHis 846
Db   3149 TCTTTGACTCAGGAGCAGAAGTTTGAACATATGTATGCTCAGAGTTTGCAGGTACCATAC 3208

Qy    847 ArgAsnThrSerLysLysAlaGlnIleIleGlySerCysTyrAlaPheSerHisAlaPhe 866
Db   3209 AGAAACTCTTTGAGGAAAGCACACATCTTTGGAATTACATTTTCCTTCACCCAGGCAATG 3268

Qy    867 IleTyrPheAlaTyrAlaAlaGlyPheArgPheGlyAlaTyrLeuIleGlnAlaGlyArg 886
```

```
             :::||||||:::||||||       ||||||||||||||||||||:::
Db   3269 ATGTATTTTTCCTATGCTGGATGTTTCCGGTTTGGAGCCTACTTGGTGGCACATAAACTA 3328

Qy    887 MetThrProGluGlyMetPheIleValPheThrAlaIleAlaTyrGlyAlaMetAlaIle 906
          |||:::   |||   :::   :::||||||:::|||:::   :::||||||||||||:::
Db   3329 ATGAGCTTTGAGGATGTTCTGTTAGTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTG 3388

Qy    907 GlyLysThrLeuValLeuAlaProGluTyrSerLysAlaLysSerGlyAlaAlaHisLeu 926
          |||:::            ||||||:::|||:::||||||||         |||||||||:::
Db   3389 GGGCAAGTGAGTTCATTTGCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATC 3448

Qy    927 PheAlaLeuLeuGluLysLysProAsnIleAspSerArgSerGlnGluGlyLysLysPro 946
                :::::||||||   |||   |||||||||   |||   ||||||         |||
Db   3449 ATCATGATCATTGAAAAAACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCG 3508

Qy    947 AspThrCysGluGlyAsnLeuGluPheArgGluValSerPhePheTyrProCysArgPro 966
          :::|||   |||||||||:::   |||   ||||||   |||   ||||||   ||||||
Db   3509 AACACATTGGAAGGAAATGTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCG 3568

Qy    967 AspValPheIleLeuArgGlyLeuSerLeuSerIleGluArgGlyLysThrValAlaPhe 986
          |||:::   :::|||:::||||||||||||   :::::::::|||:::|||:::|||
Db   3569 GACATCCCAGTCCTTCAGGGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTG 3628

Qy    987 ValGlySerSerGlyCysGlyLysSerThrSerValGlnLeuLeuGlnArgLeuTyrAsp 1006
          ||||||||||||||||||||||||||||||   ||||||||||||:::|||   ||||||
Db   3629 GTGGGCAGCAGTGGCTGTGGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGAC 3688

Qy   1007 ProValGlnGlyGlnValLeuPheAspGlyValAspAlaLysGluLeuAsnValGlnTrp 1026
          |||:::   |||:::||||||   ||||||   :::   |||   |||||||||||||||
Db   3689 CCCTTGGCAGGGAAAGTGCTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGG 3748

Qy   1027 LeuArgSerGlnIleAlaIleValProGlnGluProValLeuPheAsnCysSerIleAla 1046
          ||||||:::   :::   ||||||   |||||||||:::||||||:::||||||||||||
Db   3749 CTCCGAGCACACCTGGGCATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCT 3808

Qy   1047 GluAsnIleAlaTyrGlyAspAsnSerArgValValProLeuAspGluIleLysGluAla 1066
          ||||||||||||||||||   |||||||||||||||         :::||||||      |||
Db   3809 GAGAACATTGCCTATGGACACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCA 3868

Qy   1067 AlaAsnAlaAlaAsnIleHisSerPheIleGluGlyLeuProGluLysTyrAsnThrGln 1086
          |||      ||||||||||||:::|||||||||   ||||||      ||||||:::|||:::
Db   3869 GCAAAGGAGGCCAACATACATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAA 3928

Qy   1087 ValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGlnLysGlnArgLeuAlaIleAlaArg 1106
          ||||||   ||||||   ||||||||||||||||||||||||:::||||||||||||
Db   3929 GTAGGAGACAAAGGAACTCAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGT 3988

Qy   1107 AlaLeuLeuGlnLysProLysIleLeuLeuLeuAspGluAlaThrSerAlaLeuAspAsn 1126
          ||||||:::::::::|||   |||||||||||||||||||||||||||||||||||||||
Db   3989 GCCCTTGTTAGACAGCCTCATATTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACA 4048

Qy   1127 AspSerGluLysValValGlnHisAlaLeuAspLysAlaArgThrGlyArgThrCysLeu 1146
          :::||||||||||||||||||   ||||||||||||   |||   ||||||||||||:::
Db   4049 GAAAGTGAAAAGGTTGTCCAAGAAGCCCTGGACAAACCCAGAGAAGGCCGCACCTGCATT 4108

Qy   1147 ValValThrHisArgLeuSerAlaIleGlnAsnAlaAspLeuIleValValLeuHisAsn 1166
          |||:::   ||||||||||||   |||||||||||||||||||||||||||      |||
Db   4109 GTGATTGCTCACCGCCTGTCCACCATCCAGAATGCAGACTTAATAGTGGTGTTTCAGAAT 4168

Qy   1167 GlyLysIleLysGluGlnGlyThrHisGlnGluLeuLeuArgAsnArgAspIleTyrPhe 1186
          |||::::::|||:::   ||||||||||||:::||||||         :::   |||||||||
Db   4169 GGCAGAGTCAAGCAGCATGGCACGCATCAGCAGCTGCTGGCACAGAAAGGCATCTATTTT 4228

Qy   1187 LysLeuValAsnAlaGln 1192
             :::|||:::   |||
Db   4229 TCAATGGTCAGTGTCCAG 4246
```

RESULT 12
US-08-996-545-1
; Sequence 1, Application US/08996545
; Patent No. 5928898
; GENERAL INFORMATION:
;   APPLICANT: Skatrud, Paul L.
;   APPLICANT: de Waard, Maarten A.
;   APPLICANT: Peery, Robert B.
;   APPLICANT: Andrade, Alan C.
;   TITLE OF INVENTION: Multiple Drug Resistance Gene atrD of
;   TITLE OF INVENTION: Aspergillus Nidulans
;   NUMBER OF SEQUENCES: 3
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Eli Lilly and Company
;   STREET: Lilly Corporate Center
;   CITY: Indianapolis
;   STATE: Indiana
;   COUNTRY: U.S.
;   ZIP: 46285
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/996,545
;   FILING DATE:
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Webster, Thomas D.
;   REGISTRATION NUMBER: 39,872
;   REFERENCE/DOCKET NUMBER: X-11766
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 317-276-3334
;   TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 1:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4002 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..4002
US-08-996-545-1

```
Alignment Scores:
Pred. No.:               7.23e-247      Length:         4002
Score:                   2282.50        Matches:        494
Percent Similarity:      59.70%         Conservative:   232
Best Local Similarity:   40.62%         Mismatches:     446
Query Match:             37.97%         Indels:         45
DB:                      2              Gaps:           11
```

US-09-873-409-6 (1-1195) x US-08-996-545-1 (1-4002)

```
Qy     15 LeuProLeuMetProLeuValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu 34
          :::||||||:::|||      |||                              |||
Db    368 GTGCCATTGCTGCCGCGTCGACTTTCCAGA-----------------GGATAATGTTAT 409

Qy     35 ValGlnThrAsnThrTyrSerPhePhe------------ArgLeuThrLeuTyrTyrVal 50
          :::::::::   |||      |||:::               :::      ||||||:::|||
Db    410 ATCAAATCTCGTACGAC-GAGTTCTATGATGAATTGACCAAGAACGTACTGTACTTCGTA 468

Qy     51 GlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThr 70
             :::|||:::         :::      |||:::            :::|||   |||
Db    469 TACCTCGGTATCGGCGAGTTTGTCACTGTCTATGTTAGTACTGTTGGCTTCATCTATACC 528

Qy     71 AlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAsp 90
                   |||::::::||||||:::   :::      |||:::|||   |||:::
Db    529 GGAGAACACGCCACGCAGAAGATCCGCGAGTATTACCTTGAGTCTATCCTGCGCCAGAAC 588

Qy     91 IleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMetThr---AspIle 109
          ||||||:::||||||            ||||||:::   ||||||:::|||   |||
Db    589 ATTGGCTATTTTGATAAACTCGGTGCCGGGGAAGTGACCACCCGTATAACAGCCGATACA 648
```

```
Qy    110 AspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThr 129
Db    649 AACCTTATCCAGGATGGCATTTCGGAGAAGGTCGGTCTCACTTTGACTGCCCTGGCGACA 708

Qy    130 PheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeu 149
Db    709 TTCGTGACAGCATTCATTATCGCCTACGTCAAATACTGGAAGTTGGCTCTAATTTGCAGC 768

Qy    150 SerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeu 169
Db    769 TCAACAATTGTGGCCCTCGTTCTCACCATGGGCGGTGGTTCTCAGTTTATCATCAAGTAC 828

Qy    170 ThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSer 189
Db    829 AGCAAAAAGTCGCTTGACAGCTACGGTGCAGGCGGCACTGTTGCGGAAGAGGTCATCAGC 888

Qy    190 SerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGln 209
Db    889 TCCATCAGAAATGCCACAGCGTTTGGCACCCAAGACAAGCTTGCGAAGCAGTATGAGGTC 948

Qy    210 AsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeu 229
Db    949 CACTTAGACGAAGCTGAGAAATGGGGAACAAAGAACCAGATTGTCATGGGTTTCATGATT 1008

Qy    230 GlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSer 249
Db   1009 GGCGCCATGTTTGGCCTTATGTACTCGAACTACGGTCTTGGCTTCTGGATGGGTTCTCGT 1068

Qy    250 LeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSer 269
Db   1069 TTCCTGGTAGATGGT------GCAGTCGATGTGGGTGATATTCTCACAGTTCTCATGGCC 1122

Qy    270 ValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIle 289
Db   1123 ATCTTGATCGGATCGTTCTCCTTGGGGAACGTTAGTCCAAATGCTCAAGCATTTACAAAC 1182

Qy    290 AlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsn 309
Db   1183 GCTGTGGCCGCGGCCGCAAAGATATTTGGAACGATCGATCGCCAGTCCCCATTAGATCCA 1242

Qy    310 PheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnVal 329
Db   1243 TATTCGAACGAAGGGAAGACGCTCGACCATTTTGAGGGCCACATTGAGTTACGCAATGTC 1302

Qy    330 SerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIle 349
Db   1303 AAGCATATTTACCCATCTAGACCCGAGGTCACCGTCATGGAGGATGTTTCTCTGTCAATG 1362

Qy    350 LysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValVal 369
Db   1363 CCCGCTGGAAAAACAACCGCTTTAGTCGGCCCCTCTGGCTCTGGAAAAAGTACGGTGGTC 1422

Qy    370 GlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAsp 389
Db   1423 GGCTTGGTTGAGCGATTCTACATGCCTGTTCGCGGTACGGTTTTGCTGGATGGCCATGAC 1482

Qy    390 IleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluPro 409
Db   1483 ATCAAGGACCTCAATCTCCGCTGGCTTCGCCAACAGATCTCTTTGGTTAGCCAGGAGCCT 1542

Qy    410 ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly------------------ 423
Db   1543 GTTCTTTTTGGCACGACGATTTATAAGAATATTAGGCACGGTCTCATCGGCACAAAGTAC 1602

Qy    424 ---------ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1603 GAGAATGAATCCGAGGATAAGGTCCGGGAACTCATCGAGAACGCGGCAAAAATGGCGAAT 1662

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1663 GCTCATGACTTTATTACTGCCTTGCCTGAAGGTTATGAGACCAATGTTGGGCAGCGTGGC 1722

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   1723 TTTCTCCTTTCAGGTGGCCAGAAACAGCGCATTGCAATCGCCCGTGCCGTTGTTAGTGAC 1782

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   1783 CCAAAAATCCTGCTCCTGGATGAAGCTACTTCGGCCTTGGACACAAAATCCGAAGGCGTG 1842

Qy    501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db   1843 GTTCAAGCAGCTTTGGAGAGGGCAGCTGAAGGCCGAACTACTATTGTGATCGCTCATCGC 1902

Qy    521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db   1903 CTTTCCACGATCAAAACGGCGCACAACATTGTGGTTCTGGTCAATGGCAAAATTGCTGAA 1962

Qy    541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db   1963 CAAGGAACTCACGATGAATTGGTTGACCGCGGAGGCGCTTATCGCAAACTTGTGGAGGCT 2022

Qy    561 GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 580
Db   2023 CAACGTATCAATGAACAGAAGGAAGCTGACGCCTTGGAGGACGCCGACGCTGAGGATCTC 2082

Qy    581 ThrAsnSer------LeuProLeuHisSerValLysSerIleLysSerAspPhe------ 596
Db   2083 ACGAATGCAGATATTGCCAAAATCAAAACTGCGTCAAGCGCATCATCCGATCTCGACGGA 2142

Qy    597 ------------IleAspLys------AlaGluGluSerThrGlnSerLysGluIleSer 610
Db   2143 AAACCCACAACCATTGACCGCACGGGCACCCACAAGTCTGTTTCCAGCGCGATTCTTTCT 2202

Qy    611 ------------------LeuProGluValSerLeuLeuLysIleLeuLys--------- 621
Db   2203 AAAAGACCCCCCGAAACAACTCCGAAATACTCATTATGGACGCTGCTCAAATTTGTTGCT 2262

Qy    622 ---LeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 640
Db   2263 TCCTTCAACCGCCCTGAAATCCCGTACATGCTCATCGGTCTTGTCTTCTCAGTGTTAGCT 2322

Qy    641 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGly--- 659
Db   2323 GGTGGTGGCCAACCCACGCAAGCAGTGCTATATGCTAAAGCCATCAGCACACTCTCGCTC 2382

Qy    660 ---AsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheVal 678
Db   2383 CCAGAATCACAATATAGCAAGCTTCGACATGATGCGGATTTCTGGTCATTGATGTTCTTC 2442

Qy    679 IleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAla 698
Db   2443 GTGGTTGGTATCATTCAGTTTATCACGCAGTCAACCAATGGTGCTGCATTTGCCGTATGC 2502

Qy    699 GlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAsp 718
Db   2503 TCCGAGAGACTTATTCGTCGCGCGAGAAGCACTGCCTTTCGGACGATACTCCGTCAAGAC 2562

Qy    719 IleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIle 738
Db   2563 ATTGCTTTCTTTGACAAGGAAGAGAATAGCACCGGCGCTCTGACCTCTTTCCTGTCCACC 2622

Qy    739 AspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAla 758
Db   2623 GAGACGAAGCATCTCTCCGGTGTTAGCGGTGTGACTCTAGGCACGATCTTGATGACCTCC 2682

Qy    759 ThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeu 778
Db   2683 ACGACCCTAGGAGCGGCTATCATTATTGCCCTGGCGATTGGGTGGAAATTGGCCTTAGTT 2742

Qy    779 IleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThr 798
Db   2743 TGTATCTCGGTTGTGCCGGTTCTCCTGGCATGCGGTTTCTACCGATTCTATATGCTAGCC 2802

Qy    799 GlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAla 818
```

```
Db   2803 CAGTTTCAATCACGCTCCAAGCTTGCTTATGAGGGATCTGCAAACTTTGCTTGCGAGGCT 2862

Qy    819 LeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyr 838
                :::|||||||||:::  |||||||||||||:::          ::::::|||
Db   2863 ACATCGTCTATCCGCACAGTTGCGTCATTAACCCGGGAAAGGGATGTCTGGGAGATTTAC 2922

Qy    839 GluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySer 858
                   |||     |||  |||   :::
Db   2923 CATGCCCAGCTTGACGCACAAGGCAGGACCAGTCTAATCTCTGTCTTGAGGTCATCCCTG 2982

Qy    859 CysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGly 878
             ||||||  |||   |||   ::::::|||     |||   ||||||   :::|||
Db   2983 TTATATGCGTCGTCGCAGGCACTTGTTTTCTTCTGCGTTGCGCTCGGGTTTTGGTACGGA 3042

Qy    879 AlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAla 898
                |||:::                              |||:::  |||:::
Db   3043 GGGACACTTCTTGGTCACCACGAGTATGACATTTTCCGCTTCTTTGTTTGTTTCTCCGAG 3102

Qy    899 IleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLys 918
          |||   :::||||||   :::   |||                  ||||||:::      |||
Db   3103 ATTCTCTTTGGTGCTCAATCCGCGGGCACCGTCTTTTCCTTTGCACCAGACATGGGCAAG 3162

Qy    919 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSer 938
          ||||||:::   ||||||               |||   ::::::||||||   ||||||:::
Db   3163 GCGAAGAATGCGGCCGCCGAATTCCGACGACTGTTCGACCGAAAGCCACAAATTGATAAC 3222

Qy    939 ArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 958
             |||:::||||||:::|||   :::|||   ||||||   :::|||||||||   |||
Db   3223 TGGTCTGAAGAGGGCGAGAAGCTCGAAACGGTGGAAGGTGAAATCGAATTTAGGAACGTG 3282

Qy    959 SerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIle 978
             |||   ||||||   ||||||:::        :::|||||||||||   |||::::::
Db   3283 CACTTCAGATACCCGACCCGCCCAGAACAGCCTGTCCTGCGCGGCTTGGACCTGACCGTG 3342

Qy    979 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 998
          :::   |||:::   ||||||   ||||||   ||||||||||||||||||::::::
Db   3343 AAGCCTGGACAATATGTTGCGCTTGTCGGACCCAGCGGTTGTGGCAAGAGTACCACCATT 3402

Qy    999 GlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAsp 1018
             ||||||:::|||   ||||||   :::   |||   :::|||   ||||||   |||
Db   3403 GCATTGCTTGAGCGCTTTTACGATGCGATTGCCGGGTCCATCCTTGTTGATGGGAAGGAC 3462

Qy   1019 AlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 1038
                :::||||||:::            ||||||   :::::::::|||   |||||||||
Db   3463 ATAAGTAAACTAAATATCAACTCCTACCGCAGCTTTCTGTCACTGGTCAGCCAGGAGCCG 3522

Qy   1039 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValVal 1058
             |||:::         :::|||   |||||||||         |||            |||
Db   3523 ACACTGTACCAGGGCACCATCAAGGAAAACATCTTACTTGGTATTGTCGAAGATGACGTA 3582

Qy   1059 ProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGly 1078
          |||   :::   :::   :::|||            |||||||||:::   ||||||
Db   3583 CCGGAAGAATTCTTGATTAAGGCTTGCAAGGACGCTAATATCTACGACTTCATCATGTCG 3642

Qy   1079 LeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGln 1098
          |||||||||   :::||||||   ||||||   ||||||      ||||||||||||||||||
Db   3643 CTCCCGGAGGGCTTTAATACAGTTGTTGGCAGCAAGGGAGGCATGTTGTCTGGCGGCCAA 3702

Qy   1099 LysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAsp 1118
          |||||||||:::|||||||||||||||||||||:::      ||||||||||||||||||
Db   3703 AAGCAACGTGTGGCCATTGCCCGAGCCCTTCTTCGGGATCCCAAAATCCTTCTTCTCGAT 3762

Qy   1119 GluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLys 1138
          ||||||||||||||||||||:::::::||||||||||||||||   |||||||||
Db   3763 GAAGCGACGTCAGCCCTCGACTCCGAGTCAGAAAAGGTCGTCCAGGCGGCTTTGGATGCC 3822

Qy   1139 AlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla 1158
          |||      |||||||||   :::   |||   ||||||||||||   ||||||   |||
Db   3823 GCTGCCCGAGGCCGAACCACAATCGCCGTTGCACACCGACTCAGCACGATTCAAAAGGCG 3882

Qy   1159 AspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1178
          |||:::|||   |||         |||||||||   |||   |||||||||   ||||||
Db   3883 GACGTTATCTATGTTTTCGACCAAGGCAAGATCGTCGAAAGCGGAACGCACAGCGAACTG 3942

Qy   1179 LeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerVal 1194
          ::::::   :::         |||:::::|||||||||   ||||||:::
Db   3943 GTCCAGAAAAAGGGCCGGTACTACGAGCTGGTCAACTTGCAGAGCTTG 3990
```

RESULT 13
US-08-996-545-3
; Sequence 3, Application US/08996545
; Patent No. 5928898
; GENERAL INFORMATION:
;   APPLICANT: Skatrud, Paul L.
;   APPLICANT: de Waard, Maarten A.
;   APPLICANT: Peery, Robert B.
;   APPLICANT: Andrade, Alan C.
;   TITLE OF INVENTION: Multiple Drug Resistance Gene atrD of
;   TITLE OF INVENTION: Aspergillus Nidulans
;   NUMBER OF SEQUENCES: 3
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Eli Lilly and Company
;   STREET: Lilly Corporate Center
;   CITY: Indianapolis
;   STATE: Indiana
;   COUNTRY: U.S.
;   ZIP: 46285
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/996,545
;   FILING DATE:
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Webster, Thomas D.
;   REGISTRATION NUMBER: 39,872
;   REFERENCE/DOCKET NUMBER: X-11766
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 317-276-3334
;   TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4002 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: mRNA
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
US-08-996-545-3

Alignment Scores:

Pred. No.:	7.23e-247	Length:	4002
Score:	2282.50	Matches:	494
Percent Similarity:	59.70%	Conservative:	232
Best Local Similarity:	40.62%	Mismatches:	446
Query Match:	37.97%	Indels:	45
DB:	2	Gaps:	11

US-09-873-409-6 (1-1195) x US-08-996-545-3 (1-4002)

```
Qy     15 LeuProLeuMetProLeuValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu 34
          :::||||||:::|||      |||                              |||
Db    368 GUGCCAUUGCUGCCGCGUCGACUUUCCAGA------------------GGAUAAUGUUAU 409

Qy     35 ValGlnThrAsnThrTyrSerPhePhe------------ArgLeuThrLeuTyrTyrVal 50
          :::::::::   |||      |||:::              :::      ||||||:::|||
Db    410 AUCAAAUCUCGUACGAC-GAGUUCUAUGAUGAAUUGACCAAGAACGUACUGUACUUCGUA 468

Qy     51 GlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThr 70
             :::|||:::         :::      |||:::            :::|||   |||
Db    469 UACCUCGGUAUCGGCGAGUUUGUCACUGUCUAUGUUAGUACUGUUGGCUUCAUCUAUACC 528
```

```
Qy     71 AlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAsp 90
Db    529 GGAGAACACGCCACGCAGAAGAUCCGCGAGUAUUACCUUGAGUCUAUCCUGCGCCAGAAC 588

Qy     91 IleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMetThr---AspIle 109
Db    589 AUUGGCUAUUUUGAUAAACUCGGUGCCGGGGAAGUGACCACCCGUAUAACAGCCGAUACA 648

Qy    110 AspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThr 129
Db    649 AACCUUAUCCAGGAUGGCAUUUCGGAGAAGGUCGGUCUCACUUUGACUGCCCUGGCGACA 708

Qy    130 PheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeu 149
Db    709 UUCGUGACAGCAUUCAUUAUCGCCUACGUCAAAUACUGGAAGUUGGCUCUAAUUUGCAGC 768

Qy    150 SerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeu 169
Db    769 UCAACAAUUGUGGCCCUCGUUCUCACCAUGGGCGGUGGUUCUCAGUUUAUCAUCAAGUAC 828

Qy    170 ThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSer 189
Db    829 AGCAAAAAGUCGCUUGACAGCUACGGUGCAGGCGGCACUGUUGCGGAAGAGGUCAUCAGC 888

Qy    190 SerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGln 209
Db    889 UCCAUCAGAAAUGCCACAGCGUUUGGCACCCAAGACAAGCUUGCGAAGCAGUAUGAGGUC 948

Qy    210 AsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeu 229
Db    949 CACUUAGACGAAGCUGAGAAAUGGGGAACAAAGAACCAGAUUGUCAUGGGUUUCAUGAUU 1008

Qy    230 GlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSer 249
Db   1009 GGCGCCAUGUUUGGCCUUAUGUACUCGAACUACGGUCUUGGCUUCUGGAUGGGUUCUCGU 1068

Qy    250 LeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSer 269
Db   1069 UUCCUGGUAGAUGGU------GCAGUCGAUGUGGGUGAUAUUCUCACAGUUCUCAUGGCC 1122

Qy    270 ValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIle 289
Db   1123 AUCUUGAUCGGAUCGUUCUCCUUGGGGAACGUUAGUCCAAAUGCUCAAGCAUUUACAAAC 1182

Qy    290 AlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsn 309
Db   1183 GCUGUGGCCGCGGCCGCAAAGAUAUUUGGAACGAUCGAUCGCCAGUCCCCAUUAGAUCCA 1242

Qy    310 PheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnVal 329
Db   1243 UAUUCGAACGAAGGGAAGACGCUCGACCAUUUUGAGGGCCACAUUGAGUUACGCAAUGUC 1302

Qy    330 SerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIle 349
Db   1303 AAGCAUAUUUACCCAUCUAGACCCGAGGUCACCGUCAUGGAGGAUGUUUCUCUGUCAAUG 1362

Qy    350 LysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValVal 369
Db   1363 CCCGCUGGAAAAACAACCGCUUUAGUCGGCCCCUCUGGCUCUGGAAAAAGUACGGUGGUC 1422

Qy    370 GlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAsp 389
Db   1423 GGCUUGGUUGAGCGAUUCUACAUGCCUGUUCGCGGUACGGUUUUGCUGGAUGGCCAUGAC 1482

Qy    390 IleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluPro 409
Db   1483 AUCAAGGACCUCAAUCUCCGCUGGCUUCGCCAACAGAUCUCUUUGGUUAGCCAGGAGCCU 1542

Qy    410 ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly------------------ 423
Db   1543 GUUCUUUUUGGCACGACGAUUUAUAAGAAUAUUAGGCACGGUCUCAUCGGCACAAAGUAC 1602

Qy    424 ---------ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1603 GAGAAUGAAUCCGAGGAUAAGGUCCGGGAACUCAUCGAGAACGCGGCAAAAAUGGCGAAU 1662

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1663 GCUCAUGACUUUAUUACUGCCUUGCCUGAAGGUUAUGAGACCAAUGUUGGGCAGCGUGGC 1722

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   1723 UUUCUCCUUUCAGGUGGCCAGAAACAGCGCAUUGCAAUCGCCCGUGCCGUUGUUAGUGAC 1782

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   1783 CCAAAAAUCCUGCUCCUGGAUGAAGCUACUUCGGCCUUGGACACAAAAUCCGAAGGCGUG 1842

Qy    501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db   1843 GUUCAAGCAGCUUUGGAGAGGGCAGCUGAAGGCCGAACUACUAUUGUGAUCGCUCAUCGC 1902

Qy    521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db   1903 CUUUCCACGAUCAAAACGGCGCACAACAUUGUGGUUCUGGUCAAUGGCAAAAUUGCUGAA 1962

Qy    541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db   1963 CAAGGAACUCACGAUGAAUUGGUUGACCGCGGAGGCGCUUAUCGCAAACUUGUGGAGGCU 2022

Qy    561 GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 580
Db   2023 CAACGUAUCAAUGAACAGAAGGAAGCUGACGCCUUGGAGGACGCCGACGCUGAGGAUCUC 2082

Qy    581 ThrAsnSer------LeuProLeuHisSerValLysSerIleLysSerAspPhe------ 596
Db   2083 ACGAAUGCAGAUAUUGCCAAAAUCAAAACUGCGUCAAGCGCAUCAUCCGAUCUCGACGGA 2142

Qy    597 ------------IleAspLys------AlaGluGluSerThrGlnSerLysGluIleSer 610
Db   2143 AAACCCACAACCAUUGACCGCACGGGCACCCACAAGUCUGUUUCCAGCGCGAUUCUUUCU 2202

Qy    611 ------------------LeuProGluValSerLeuLeuLysIleLeuLys--------- 621
Db   2203 AAAAGACCCCCCGAAACAACUCCGAAAUACUCAUUAUGGACGCUGCUCAAAUUUGUUGCU 2262

Qy    622 ---LeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 640
Db   2263 UCCUUCAACCGCCCUGAAAUCCCGUACAUGCUCAUCGGUCUUGUCUUCUCAGUGUUAGCU 2322

Qy    641 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGly--- 659
Db   2323 GGUGGUGGCCAACCCACGCAAGCAGUGCUAUAUGCUAAAGCCAUCAGCACACUCUCGCUC 2382

Qy    660 ---AsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheVal 678
Db   2383 CCAGAAUCACAAUAUAGCAAGCUUCGACAUGAUGCGGAUUUCUGGUCAUUGAUGUUCUUC 2442

Qy    679 IleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAla 698
Db   2443 GUGGUUGGUAUCAUUCAGUUUAUCACGCAGUCAACCAAUGGUGCUGCAUUUGCCGUAUGC 2502

Qy    699 GlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAsp 718
Db   2503 UCCGAGAGACUUAUUCGUCGCGCGAGAAGCACUGCCUUUCGGACGAUACUCCGUCAAGAC 2562

Qy    719 IleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIle 738
Db   2563 AUUGCUUUCUUUGACAAGGAAGAGAAUAGCACCGGCGCUCUGACCUCUUUCCUGUCCACC 2622

Qy    739 AspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAla 758
Db   2623 GAGACGAAGCAUCUCUCCGGUGUUAGCGGUGUGACUCUAGGCACGAUCUUGAUGACCUCC 2682

Qy    759 ThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeu 778
```

```
Db   2683 ACGACCCUAGGAGCGGCUAUCAUUAUUGCCCUGGCGAUUGGGUGGAAAUUGGCCUUAGUU 2742

Qy    779 IleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThr 798
Db   2743 UGUAUCUCGGUUGUGCCGGUUCUCCUGGCAUGCGGUUUCUACCGAUUCUAUAUGCUAGCC 2802

Qy    799 GlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAla 818
Db   2803 CAGUUUCAAUCACGCUCCAAGCUUGCUUAUGAGGGAUCUGCAAACUUUGCUUGCGAGGCU 2862

Qy    819 LeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyr 838
Db   2863 ACAUCGUCUAUCCGCACAGUUGCGUCAUUAACCCGGGAAAGGGAUGUCUGGGAGAUUUAC 2922

Qy    839 GluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySer 858
Db   2923 CAUGCCCAGCUUGACGCACAAGGCAGGACCAGUCUAAUCUCUGUCUUGAGGUCAUCCCUG 2982

Qy    859 CysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGly 878
Db   2983 UUAUAUGCGUCGUCGCAGGCACUUGUUUUCUUCUGCGUUGCGCUCGGGUUUUGGUACGGA 3042

Qy    879 AlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAla 898
Db   3043 GGGACACUUCUUGGUCACCACGAGUAUGACAUUUUCCGCUUCUUUGUUUGUUUCUCCGAG 3102

Qy    899 IleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLys 918
Db   3103 AUUCUCUUUGGUGCUCAAUCCGCGGGCACCGUCUUUUCCUUUGCACCAGACAUGGGCAAG 3162

Qy    919 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSer 938
Db   3163 GCGAAGAAUGCGGCCGCCGAAUUCCGACGACUGUUCGACCGAAAGCCACAAAUUGAUAAC 3222

Qy    939 ArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 958
Db   3223 UGGUCUGAAGAGGGCGAGAAGCUCGAAACGGUGGAAGGUGAAAUCGAAUUUAGGAACGUG 3282

Qy    959 SerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIle 978
Db   3283 CACUUCAGAUACCCGACCCGCCCAGAACAGCCUGUCCUGCGCGGCUUGGACCUGACCGUG 3342

Qy    979 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 998
Db   3343 AAGCCUGGACAAUAUGUUGCGCUUGUCGGACCCAGCGGUUGUGGCAAGAGUACCACCAUU 3402

Qy    999 GlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAsp 1018
Db   3403 GCAUUGCUUGAGCGCUUUUACGAUGCGAUUGCCGGGUCCAUCCUUGUUGAUGGGAAGGAC 3462

Qy   1019 AlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 1038
Db   3463 AUAAGUAAACUAAAUAUCAACUCCUACCGCAGCUUUCUGUCACUGGUCAGCCAGGAGCCG 3522

Qy   1039 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValVal 1058
Db   3523 ACACUGUACCAGGGCACCAUCAAGGAAAACAUCUUACUUGGUAUUGUCGAAGAUGACGUA 3582

Qy   1059 ProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGly 1078
Db   3583 CCGGAAGAAUUCUUGAUUAAGGCUUGCAAGGACGCUAAUAUCUACGACUUCAUCAUGUCG 3642

Qy   1079 LeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGln 1098
Db   3643 CUCCCGGAGGGCUUUAAUACAGUUGUUGGCAGCAAGGGAGGCAUGUUGUCUGGCGGCCAA 3702

Qy   1099 LysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAsp 1118
Db   3703 AAGCAACGUGUGGCCAUUGCCCGAGCCCUUCUUCGGGAUCCCAAAAUCCUUCUUCUCGAU 3762

Qy   1119 GluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLys 1138
Db   3763 GAAGCGACGUCAGCCCUCGACUCCGAGUCAGAAAAGGUCGUCCAGGCGGCUUUGGAUGCC 3822

Qy   1139 AlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla 1158
Db   3823 GCUGCCCGAGGCCGAACCACAAUCGCCGUUGCACACCGACUCAGCACGAUUCAAAAGGCG 3882

Qy   1159 AspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1178
Db   3883 GACGUUAUCUAUGUUUUCGACCAAGGCAAGAUCGUCGAAAGCGGAACGCACAGCGAACUG 3942

Qy   1179 LeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerVal 1194
Db   3943 GUCCAGAAAAAGGGCCGGUACUACGAGCUGGUCAACUUGCAGAGCUUG 3990
```

```
RESULT 14
US-09-328-320-1
; Sequence 1, Application US/09328320
; Patent No. 6228615
;  GENERAL INFORMATION:
;    APPLICANT:  Skatrud, Paul L.
;    APPLICANT:  de Waard, Maarten A.
;    APPLICANT:  Peery, Robert B.
;    APPLICANT:  Andrade, Alan C.
;    TITLE OF INVENTION:  Multiple Drug Resistance Gene atrD of
;    TITLE OF INVENTION:  Aspergillus Nidulans
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  U.S.
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/328,320
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/996,545
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Webster, Thomas D.
;      REGISTRATION NUMBER:  39,872
;      REFERENCE/DOCKET NUMBER:  X-11766
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-3334
;      TELEFAX:  317-276-2763
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4002 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..4002
US-09-328-320-1

Alignment Scores:
Pred. No.:               7.23e-247     Length:         4002
Score:                   2282.50       Matches:        494
Percent Similarity:      59.70%        Conservative:   232
Best Local Similarity:   40.62%        Mismatches:     446
Query Match:             37.97%        Indels:         45
DB:                      4             Gaps:           11
```

```
US-09-873-409-6 (1-1195) x US-09-328-320-1 (1-4002)

Qy     15 LeuProLeuMetProLeuValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu 34
Db    368 GTGCCATTGCTGCCGCGTCGACTTTCCAGA-----------------GGATAATGTTAT 409

Qy     35 ValGlnThrAsnThrTyrSerPhePhe------------ArgLeuThrLeuTyrTyrVal 50
Db    410 ATCAAATCTCGTACGAC-GAGTTCTATGATGAATTGACCAAGAACGTACTGTACTTCGTA 468

Qy     51 GlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThr 70
Db    469 TACCTCGGTATCGGCGAGTTTGTCACTGTCTATGTTAGTACTGTTGGCTTCATCTATACC 528

Qy     71 AlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAsp 90
Db    529 GGAGAACACGCCACGCAGAAGATCCGCGAGTATTACCTTGAGTCTATCCTGCGCCAGAAC 588

Qy     91 IleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMetThr---AspIle 109
Db    589 ATTGGCTATTTGATAAACTCGGTGCCGGGGAAGTGACCACCCGTATAACAGCCGATACA 648

Qy    110 AspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThr 129
Db    649 AACCTTATCCAGGATGGCATTTCGGAGAAGGTCGGTCTCACTTTGACTGCCCTGGCGACA 708

Qy    130 PheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeu 149
Db    709 TTCGTGACAGCATTCATTATCGCCTACGTCAAATACTGGAAGTTGGCTCTAATTTGCAGC 768

Qy    150 SerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeu 169
Db    769 TCAACAATTGTGGCCCTCGTTCTCACCATGGGCGGTGGTTCTCAGTTTATCATCAAGTAC 828

Qy    170 ThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSer 189
Db    829 AGCAAAAAGTCGCTTGACAGCTACGGTGCAGGCGGCACTGTTGCGGAAGAGGTCATCAGC 888

Qy    190 SerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGln 209
Db    889 TCCATCAGAAATGCCACAGCGTTTGGCACCCAAGACAAGCTTGCGAAGCAGTATGAGGTC 948

Qy    210 AsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeu 229
Db    949 CACTTAGACGAAGCTGAGAAATGGGGAACAAAGAACCAGATTGTCATGGGTTTCATGATT 1008

Qy    230 GlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSer 249
Db   1009 GGCGCCATGTTTGGCCTTATGTACTCGAACTACGGTCTTGGCTTCTGGATGGGTTCTCGT 1068

Qy    250 LeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSer 269
Db   1069 TTCCTGGTAGATGGT------GCAGTCGATGTGGGTGATATTCTCACAGTTCTCATGGCC 1122

Qy    270 ValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIle 289
Db   1123 ATCTTGATCGGATCGTTCTCCTTGGGGAACGTTAGTCCAAATGCTCAAGCATTTACAAAC 1182

Qy    290 AlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsn 309
Db   1183 GCTGTGGCCGCGGCCGCAAAGATATTTGGAACGATCGATCGCCAGTCCCCATTAGATCCA 1242

Qy    310 PheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnVal 329
Db   1243 TATTCGAACGAAGGGAAGACGCTCGACCATTTTGAGGGCCACATTGAGTTACGCAATGTC 1302

Qy    330 SerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIle 349
Db   1303 AAGCATATTTACCCATCTAGACCCGAGGTCACCGTCATGGAGGATGTTTCTCTGTCAATG 1362

Qy    350 LysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValVal 369
Db   1363 CCCGCTGGAAAAACAACCGCTTTAGTCGGCCCCTCTGGCTCTGGAAAAAGTACGGTGGTC 1422

Qy    370 GlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAsp 389
Db   1423 GGCTTGGTTGAGCGATTCTACATGCCTGTTCGCGGTACGGTTTTGCTGGATGGCCATGAC 1482

Qy    390 IleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluPro 409
Db   1483 ATCAAGGACCTCAATCTCCGCTGGCTTCGCCAACAGATCTCTTTGGTTAGCCAGGAGCCT 1542

Qy    410 ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly------------------ 423
Db   1543 GTTCTTTTGGCACGACGATTTATAAGAATATTAGGCACGGTCTCATCGGCACAAAGTAC 1602

Qy    424 ---------ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1603 GAGAATGAATCCGAGGATAAGGTCCGGGAACTCATCGAGAACGCGGCAAAAATGGCGAAT 1662

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1663 GCTCATGACTTTATTACTGCCTTGCCTGAAGGTTATGAGACCAATGTTGGGCAGCGTGGC 1722

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   1723 TTTCTCCTTTCAGGTGGCCAGAAACAGCGCATTGCAATCGCCCGTGCCGTTGTTAGTGAC 1782

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   1783 CCAAAAATCCTGCTCCTGGATGAAGCTACTTCGGCCTTGGACACAAAATCCGAAGGCGTG 1842

Qy    501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db   1843 GTTCAAGCAGCTTTGGAGAGGGCAGCTGAAGGCCGAACTACTATTGTGATCGCTCATCGC 1902

Qy    521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db   1903 CTTTCCACGATCAAAACGGCGCACAACATTGTGGTTCTGGTCAATGGCAAAATTGCTGAA 1962

Qy    541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db   1963 CAAGGAACTCACGATGAATTGGTTGACCGCGGAGGCGCTTATCGCAAACTTGTGGAGGCT 2022

Qy    561 GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 580
Db   2023 CAACGTATCAATGAACAGAAGGAAGCTGACGCCTTGGAGGACGCCGACGCTGAGGATCTC 2082

Qy    581 ThrAsnSer------LeuProLeuHisSerValLysSerIleLysSerAspPhe------ 596
Db   2083 ACGAATGCAGATATTGCCAAAATCAAAACTGCGTCAAGCGCATCATCCGATCTCGACGGA 2142

Qy    597 ------------IleAspLys------AlaGluGluSerThrGlnSerLysGluIleSer 610
Db   2143 AAACCCACAACCATTGACCGCACGGGCACCCACAAGTCTGTTTCCAGCGCGATTCTTTCT 2202

Qy    611 ------------------LeuProGluValSerLeuLeuLysIleLeuLys--------- 621
Db   2203 AAAAGACCCCCCGAAACAACTCCGAAATACTCATTATGGACGCTGCTCAAATTTGTTGCT 2262

Qy    622 ---LeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 640
Db   2263 TCCTTCAACCGCCCTGAAATCCCGTACATGCTCATCGGTCTTGTCTTCTCAGTGTTAGCT 2322

Qy    641 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGly--- 659
Db   2323 GGTGGTGGCCAACCCACGCAAGCAGTGCTATATGCTAAAGCCATCAGCACACTCTCGCTC 2382

Qy    660 ---AsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheVal 678
Db   2383 CCAGAATCACAATATAGCAAGCTTCGACATGATGCGGATTTCTGGTCATTGATGTTCTTC 2442

Qy    679 IleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAla 698
Db   2443 GTGGTTGGTATCATTCAGTTTATCACGCAGTCAACCAATGGTGCTGCATTTGCCGTATGC 2502
```

```
Qy     699 GlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAsp 718
Db    2503 TCCGAGAGACTTATTCGTCGCGCGAGAAGCACTGCCTTTCGGACGATACTCCGTCAAGAC 2562

Qy     719 IleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIle 738
Db    2563 ATTGCTTTCTTTGACAAGGAAGAGAATAGCACCGGCGCTCTGACCTCTTTCCTGTCCACC 2622

Qy     739 AspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAla 758
Db    2623 GAGACGAAGCATCTCTCCGGTGTTAGCGGTGTGACTCTAGGCACGATCTTGATGACCTCC 2682

Qy     759 ThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeu 778
Db    2683 ACGACCCTAGGAGCGGCTATCATTATTGCCCTGGCGATTGGGTGGAAATTGGCCTTAGTT 2742

Qy     779 IleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThr 798
Db    2743 TGTATCTCGGTTGTGCCGGTTCTCCTGGCATGCGGTTTCTACCGATTCTATATGCTAGCC 2802

Qy     799 GlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAla 818
Db    2803 CAGTTTCAATCACGCTCCAAGCTTGCTTATGAGGGATCTGCAAACTTTGCTTGCGAGGCT 2862

Qy     819 LeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyr 838
Db    2863 ACATCGTCTATCCGCACAGTTGCGTCATTAACCCGGGAAAGGGATGTCTGGGAGATTTAC 2922

Qy     839 GluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySer 858
Db    2923 CATGCCCAGCTTGACGCACAAGGCAGGACCAGTCTAATCTCTGTCTTGAGGTCATCCCTG 2982

Qy     859 CysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGly 878
Db    2983 TTATATGCGTCGTCGCAGGCACTTGTTTTCTTCTGCGTTGCGCTCGGGTTTTGGTACGGA 3042

Qy     879 AlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAla 898
Db    3043 GGGACACTTCTTGGTCACCACGAGTATGACATTTTCCGCTTCTTTGTTTGTTTCTCCGAG 3102

Qy     899 IleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLys 918
Db    3103 ATTCTCTTTGGTGCTCAATCCGCGGGCACCGTCTTTTCCTTTGCACCAGACATGGGCAAG 3162

Qy     919 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSer 938
Db    3163 GCGAAGAATGCGGCCGCCGAATTCCGACGACTGTTCGACCGAAAGCCACAAATTGATAAC 3222

Qy     939 ArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 958
Db    3223 TGGTCTGAAGAGGGCGAGAAGCTCGAAACGGTGGAAGGTGAAATCGAATTTAGGAACGTG 3282

Qy     959 SerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIle 978
Db    3283 CACTTCAGATACCCGACCCGCCCAGAACAGCCTGTCCTGCGCGGCTTGGACCTGACCGTG 3342

Qy     979 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 998
Db    3343 AAGCCTGGACAATATGTTGCGCTTGTCGGACCCAGCGGTTGTGGCAAGAGTACCACCATT 3402

Qy     999 GlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAsp 1018
Db    3403 GCATTGCTTGAGCGCTTTTACGATGCGATTGCCGGGTCCATCCTTGTTGATGGGAAGGAC 3462

Qy    1019 AlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 1038
Db    3463 ATAAGTAAACTAAATATCAACTCCTACCGCAGCTTTCTGTCACTGGTCAGCCAGGAGCCG 3522

Qy    1039 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValVal 1058
Db    3523 ACACTGTACCAGGGCACCATCAAGGAAAACATCTTACTTGGTATTGTCGAAGATGACGTA 3582

Qy    1059 ProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGly 1078
Db    3583 CCGGAAGAATTCTTGATTAAGGCTTGCAAGGACGCTAATATCTACGACTTCATCATGTCG 3642

Qy    1079 LeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGln 1098
Db    3643 CTCCCGGAGGGCTTTAATACAGTTGTTGGCAGCAAGGGAGGCATGTTGTCTGGCGGCCAA 3702

Qy    1099 LysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAsp 1118
Db    3703 AAGCAACGTGTGGCCATTGCCCGAGCCCTTCTTCGGGATCCCAAAATCCTTCTTCTCGAT 3762

Qy    1119 GluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLys 1138
Db    3763 GAAGCGACGTCAGCCCTCGACTCCGAGTCAGAAAAGGTCGTCCAGGCGGCTTTGGATGCC 3822

Qy    1139 AlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla 1158
Db    3823 GCTGCCCGAGGCCGAACCACAATCGCCGTTGCACACCGACTCAGCACGATTCAAAAGGCG 3882

Qy    1159 AspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1178
Db    3883 GACGTTATCTATGTTTTCGACCAAGGCAAGATCGTCGAAAGCGGAACGCACAGCGAACTG 3942

Qy    1179 LeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerVal 1194
Db    3943 GTCCAGAAAAAGGGCCGGTACTACGAGCTGGTCAACTTGCAGAGCTTG 3990
```

RESULT 15
US-09-328-320-3
; Sequence 3, Application US/09328320
; Patent No. 6228615
; GENERAL INFORMATION:
;   APPLICANT: Skatrud, Paul L.
;   APPLICANT: de Waard, Maarten A.
;   APPLICANT: Peery, Robert B.
;   APPLICANT: Andrade, Alan C.
;   TITLE OF INVENTION: Multiple Drug Resistance Gene atrD of
;   TITLE OF INVENTION: Aspergillus Nidulans
;   NUMBER OF SEQUENCES: 3
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Eli Lilly and Company
;   STREET: Lilly Corporate Center
;   CITY: Indianapolis
;   STATE: Indiana
;   COUNTRY: U.S.
;   ZIP: 46285
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/09/328,320
;   FILING DATE:
;   CLASSIFICATION:
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 08/996,545
;   FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Webster, Thomas D.
;   REGISTRATION NUMBER: 39,872
;   REFERENCE/DOCKET NUMBER: X-11766
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 317-276-3334
;   TELEFAX: 317-276-2763
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4002 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: mRNA
;   HYPOTHETICAL: NO

```
;    ANTI-SENSE:  NO
US-09-328-320-3

Alignment Scores:
Pred. No.:              7.23e-247      Length:        4002
Score:                  2282.50        Matches:       494
Percent Similarity:     59.70%         Conservative:  232
Best Local Similarity:  40.62%         Mismatches:    446
Query Match:            37.97%         Indels:        45
DB:                     4              Gaps:          11

US-09-873-409-6 (1-1195) x US-09-328-320-3 (1-4002)

Qy     15 LeuProLeuMetProLeuValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu 34
Db    368 GUGCCAUUGCUGCCGCGUCGACUUUCCAGA-----------------GGAUAAUGUUAU 409

Qy     35 ValGlnThrAsnThrTyrSerPhePhe------------ArgLeuThrLeuTyrTyrVal 50
Db    410 AUCAAAUCUCGUACGAC-GAGUUCUAUGAUGAAUUGACCAAGAACGUACUGUACUUCGUA 468

Qy     51 GlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThr 70
Db    469 UACCUCGGUAUCGGCGAGUUUGUCACUGUCUAUGUUAGUACUGUUGGCUUCAUCUAUACC 528

Qy     71 AlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAsp 90
Db    529 GGAGAACACGCCACGCAGAAGAUCCGCGAGUAUUACCUUGAGUCUAUCCUGCGCCAGAAC 588

Qy     91 IleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMetThr---AspIle 109
Db    589 AUUGGCUAUUUUGAUAAACUCGGUGCCGGGGAAGUGACCACCCGUAUAACAGCCGAUACA 648

Qy    110 AspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThr 129
Db    649 AACCUUAUCCAGGAUGGCAUUUCGGAGAAGGUCGGUCUCACUUUGACUGCCCUGGCGACA 708

Qy    130 PheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeu 149
Db    709 UUCGUGACAGCAUUCAUUAUCGCCUACGUCAAAUACUGGAAGUUGGCUCUAAUUUGCAGC 768

Qy    150 SerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeu 169
Db    769 UCAACAAUUGUGGCCCUCGUUCUCACCAUGGGCGGUGGUUCUCAGUUUAUCAUCAAGUAC 828

Qy    170 ThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSer 189
Db    829 AGCAAAAAGUCGCUUGACAGCUACGGUGCAGGCGGCACUGUUGCGGAAGAGGUCAUCAGC 888

Qy    190 SerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGln 209
Db    889 UCCAUCAGAAAUGCCACAGCGUUUGGCACCCAAGACAAGCUUGCGAAGCAGUAUGAGGUC 948

Qy    210 AsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeu 229
Db    949 CACUUAGACGAAGCUGAGAAAUGGGGAACAAAGAACCAGAUUGUCAUGGGUUUCAUGAUU 1008

Qy    230 GlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSer 249
Db   1009 GGCGCCAUGUUUGGCCUUAUGUACUCGAACUACGGUCUUGGCUUCUGGAUGGGUUCUCGU 1068

Qy    250 LeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSer 269
Db   1069 UUCCUGGUAGAUGGU------GCAGUCGAUGUGGGUGAUAUUCUCACAGUUCUCAUGGCC 1122

Qy    270 ValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIle 289
Db   1123 AUCUUGAUCGGAUCGUUCUCCUUGGGGAACGUUAGUCCAAAUGCUCAAGCAUUUACAAAC 1182

Qy    290 AlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsn 309
Db   1183 GCUGUGGCCGCGGCCGCAAAGAUAUUUGGAACGAUCGAUCGCCAGUCCCCAUUAGAUCCA 1242

Qy    310 PheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnVal 329
Db   1243 UAUUCGAACGAAGGGAAGACGCUCGACCAUUUUGAGGGCCACAUUGAGUUACGCAAUGUC 1302

Qy    330 SerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIle 349
Db   1303 AAGCAUAUUUACCCAUCUAGACCCGAGGUCACCGUCAUGGAGGAUGUUUCUCUGUCAAUG 1362

Qy    350 LysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValVal 369
Db   1363 CCCGCUGGAAAAACAACCGCUUUAGUCGGCCCCUCUGGCUCUGGAAAAAGUACGGUGGUC 1422

Qy    370 GlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAsp 389
Db   1423 GGCUUGGUUGAGCGAUUCUACAUGCCUGUUCGCGGUACGGUUUUGCUGGAUGGCCAUGAC 1482

Qy    390 IleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluPro 409
Db   1483 AUCAAGGACCUCAAUCUCCGCUGGCUUCGCCAACAGAUCUCUUUGGUUAGCCAGGAGCCU 1542

Qy    410 ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly------------------ 423
Db   1543 GUUCUUUUUGGCACGACGAUUUAUAAGAAUAUUAGCACGGUCUCAUCGGCACAAAGUAC 1602

Qy    424 ---------ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1603 GAGAAUGAAUCCGAGGAUAAGGUCCGGGAACUCAUCGAGAACGCGGCAAAAAUGGCGAAU 1662

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1663 GCUCAUGACUUUAUUACUGCCUUGCCUGAAGGUUAUGAGACCAAUGUUGGGCAGCGUGGC 1722

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   1723 UUUCUCCUUUCAGGUGGCCAGAAACAGCGCAUUGCAAUCGCCCGUGCCGUUGUUAGUGAC 1782

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   1783 CCAAAAAUCCUGCUCCUGGAUGAAGCUACUUCGGCCUUGGACACAAAAUCCGAAGGCGUG 1842

Qy    501 ValGlnAlaAlaLeuGluLysAlaSerLysGlyArgThrThrIleValValAlaHisArg 520
Db   1843 GUUCAAGCAGCUUUGGAGAGGGCAGCUGAAGGCCGAACUACUAUUGUGAUCGCUCAUCGC 1902

Qy    521 LeuSerThrIleArgSerAlaAspLeuIleValThrLeuLysAspGlyMetLeuAlaGlu 540
Db   1903 CUUUCCACGAUCAAAACGGCGCACAACAUUGUGGUUCUGGUCAAUGGCAAAAUUGCUGAA 1962

Qy    541 LysGlyAlaHisAlaGluLeuMetAlaLysArgGlyLeuTyrTyrSerLeuValMetSer 560
Db   1963 CAAGGAACUCACGAUGAAUUGGUUGACCGCGGAGGCGCUUAUCGCAAACUUGUGGAGGCU 2022

Qy    561 GlnAspIleLysLysAlaAspGluGlnMetGluSerMetThrTyrSerThrGluArgLys 580
Db   2023 CAACGUAUCAAUGAACAGAAGGAAGCUGACGCCUUGGAGGACGCCGACGCUGAGGAUCUC 2082

Qy    581 ThrAsnSer------LeuProLeuHisSerValLysSerIleLysSerAspPhe------ 596
Db   2083 ACGAAUGCAGAUAUUGCCAAAAUCAAAACUGCGUCAAGCGCAUCAUCCGAUCUCGACGGA 2142

Qy    597 ------------IleAspLys------AlaGluGluSerThrGlnSerLysGluIleSer 610
Db   2143 AAACCCACAACCAUUGACCGCACGGGCACCCACAAGUCUGUUUCCAGCGCGAUUCUUUCU 2202

Qy    611 ------------------LeuProGluValSerLeuLeuLysIleLeuLys--------- 621
Db   2203 AAAAGACCCCCCGAAACAACUCCGAAAUACUCAUUAUGGACGCUGCUCAAAUUUGUUGCU 2262

Qy    622 ---LeuAsnLysProGluTrpProPheValValLeuGlyThrLeuAlaSerValLeuAsn 640
Db   2263 UCCUUCAACCGCCCUGAAAUCCCGUACAUGCUCAUCGGUCUUGUCUUCUCAGUGUUAGCU 2322

Qy    641 GlyThrValHisProValPheSerIleIlePheAlaLysIleIleThrMetPheGly--- 659
```

```
Db   2323 GGUGGUGGCCAACCCACGCAAGCAGUGCUAUAUGCUAAAGCCAUCAGCACACUCUCGCUC 2382

Qy    660 ---AsnAsnAspLysThrThrLeuLysHisAspAlaGluIleTyrSerMetIlePheVal 678
Db   2383 CCAGAAUCACAAUAUAGCAAGCUUCGACAUGAUGCGGAUUUCUGGUCAUUGAUGUUCUUC 2442

Qy    679 IleLeuGlyValIleCysPheValSerTyrPheMetGlnGlyLeuPheTyrGlyArgAla 698
Db   2443 GUGGUUGGUAUCAUUCAGUUUAUCACGCAGUCAACCAAUGGUGCUGCAUUUGCCGUAUGC 2502

Qy    699 GlyGluIleLeuThrMetArgLeuArgHisLeuAlaPheLysAlaMetLeuTyrGlnAsp 718
Db   2503 UCCGAGAGACUUAUUCGUCGCGCGAGAAGCACUGCCUUUCGGACGAUACUCCGUCAAGAC 2562

Qy    719 IleAlaTrpPheAspGluLysGluAsnSerThrGlyGlyLeuThrThrIleLeuAlaIle 738
Db   2563 AUUGCUUUCUUUGACAAGGAAGAGAAUAGCACCGGCGCUCUGACCUCUUUCCUGUCCACC 2622

Qy    739 AspIleAlaGlnIleGlnGlyAlaThrGlySerArgIleGlyValLeuThrGlnAsnAla 758
Db   2623 GAGACGAAGCAUCUCUCCGGUGUUAGCGGUGUGACUCUAGGCACGAUCUUGAUGACCUCC 2682

Qy    759 ThrAsnMetGlyLeuSerValIleIleSerPheIleTyrGlyTrpGluMetThrPheLeu 778
Db   2683 ACGACCCUAGGAGCGGCUAUCAUUAUUGCCCUGGCGAUUGGGUGGAAAUUGGCCUUAGUU 2742

Qy    779 IleLeuSerIleAlaProValLeuAlaValThrGlyMetIleGluThrAlaAlaMetThr 798
Db   2743 UGUAUCUCGGUUGUGCCGGUUCUCCUGGCAUGCGGUUUCUACCGAUUCUAUAUGCUAGCC 2802

Qy    799 GlyPheAlaAsnLysAspLysGlnGluLeuLysHisAlaGlyLysIleAlaThrGluAla 818
Db   2803 CAGUUUCAAUCACGCUCCAAGCUUGCUUAUGAGGGAUCUGCAAACUUUGCUUGCGAGGCU 2862

Qy    819 LeuGluAsnIleArgThrIleValSerLeuThrArgGluLysAlaPheGluGlnMetTyr 838
Db   2863 ACAUCGUCUAUCCGCACAGUUGCGUCAUUAACCCGGGAAAGGGAUGUCUGGGAGAUUUAC 2922

Qy    839 GluGluMetLeuGlnThrGlnHisArgAsnThrSerLysLysAlaGlnIleIleGlySer 858
Db   2923 CAUGCCCAGCUUGACGCACAAGGCAGGACCAGUCUAAUCUCUGUCUUGAGGUCAUCCCUG 2982

Qy    859 CysTyrAlaPheSerHisAlaPheIleTyrPheAlaTyrAlaAlaGlyPheArgPheGly 878
Db   2983 UUAUAUGCGUCGUCGCAGGCACUUGUUUUCUUCUGCGUUGCGCUCGGGUUUUGGUACGGA 3042

Qy    879 AlaTyrLeuIleGlnAlaGlyArgMetThrProGluGlyMetPheIleValPheThrAla 898
Db   3043 GGGACACUUCUUGGUCACCACGAGUAUGACAUUUUCCGCUUCUUUGUUUGUUUCUCCGAG 3102

Qy    899 IleAlaTyrGlyAlaMetAlaIleGlyLysThrLeuValLeuAlaProGluTyrSerLys 918
Db   3103 AUUCUCUUUGGUGCUCAAUCCGCGGGCACCGUCUUUUCCUUUGCACCAGACAUGGGCAAG 3162

Qy    919 AlaLysSerGlyAlaAlaHisLeuPheAlaLeuLeuGluLysLysProAsnIleAspSer 938
Db   3163 GCGAAGAAUGCGGCCGCCGAAUUCCGACGACUGUUCGACCGAAAGCCACAAAUUGAUAAC 3222

Qy    939 ArgSerGlnGluGlyLysLysProAspThrCysGluGlyAsnLeuGluPheArgGluVal 958
Db   3223 UGGUCUGAAGAGGGCGAGAAGCUCGAAACGGUGGAAGGUGAAAUCGAAUUUAGGAACGUG 3282

Qy    959 SerPhePheTyrProCysArgProAspValPheIleLeuArgGlyLeuSerLeuSerIle 978
Db   3283 CACUUCAGAUACCCGACCCGCCCAGAACAGCCUGUCCUGCGCGGCUUGGACCUGACCGUG 3342

Qy    979 GluArgGlyLysThrValAlaPheValGlySerSerGlyCysGlyLysSerThrSerVal 998
Db   3343 AAGCCUGGACAAUAUGUUGCGCUUGUCGGACCCAGCGGUUGUGGCAAGAGUACCACCAUU 3402

Qy    999 GlnLeuLeuGlnArgLeuTyrAspProValGlnGlyGlnValLeuPheAspGlyValAsp 1018
Db   3403 GCAUUGCUUGAGCGCUUUUACGAUGCGAUUGCCGGGUCCAUCCUUGUUGAUGGGAAGGAC 3462

Qy   1019 AlaLysGluLeuAsnValGlnTrpLeuArgSerGlnIleAlaIleValProGlnGluPro 1038
Db   3463 AUAAGUAAACUAAAUAUCAACUCCUACCGCAGCUUUCUGUCACUGGUCAGCCAGGAGCCG 3522

Qy   1039 ValLeuPheAsnCysSerIleAlaGluAsnIleAlaTyrGlyAspAsnSerArgValVal 1058
Db   3523 ACACUGUACCAGGGCACCAUCAAGGAAAACAUCUUACUUGGUAUUGUCGAAGAUGACGUA 3582

Qy   1059 ProLeuAspGluIleLysGluAlaAlaAsnAlaAlaAsnIleHisSerPheIleGluGly 1078
Db   3583 CCGGAAGAAUUCUUGAUUAAGGCUUGCAAGGACGCUAAUAUCUACGACUUCAUCAUGUCG 3642

Qy   1079 LeuProGluLysTyrAsnThrGlnValGlyLeuLysGlyAlaGlnLeuSerGlyGlyGln 1098
Db   3643 CUCCCGGAGGGCUUUAAUACAGUUGUUGGCAGCAAGGGAGGCAUGUUGUCUGGCGGCCAA 3702

Qy   1099 LysGlnArgLeuAlaIleAlaArgAlaLeuLeuGlnLysProLysIleLeuLeuLeuAsp 1118
Db   3703 AAGCAACGUGUGGCCAUUGCCCGAGCCCUUCUUCGGGAUCCCAAAAUCCUUCUUCUCGAU 3762

Qy   1119 GluAlaThrSerAlaLeuAspAsnAspSerGluLysValValGlnHisAlaLeuAspLys 1138
Db   3763 GAAGCGACGUCAGCCCUCGACUCCGAGUCAGAAAAGGUCGUCCAGGCGGCUUUGGAUGCC 3822

Qy   1139 AlaArgThrGlyArgThrCysLeuValValThrHisArgLeuSerAlaIleGlnAsnAla 1158
Db   3823 GCUGCCCGAGGCCGAACCACAAUCGCCGUUGCACACCGACUCAGCACGAUUCAAAAGGCG 3882

Qy   1159 AspLeuIleValValLeuHisAsnGlyLysIleLysGluGlnGlyThrHisGlnGluLeu 1178
Db   3883 GACGUUAUCUAUGUUUUCGACCAAGGCAAGAUCGUCGAAAGCGGAACGCACAGCGAACUG 3942

Qy   1179 LeuArgAsnArgAspIleTyrPheLysLeuValAsnAlaGlnSerVal 1194
Db   3943 GUCCAGAAAAAGGGCCGGUACUACGAGCUGGUCAACUUGCAGAGCUUG 3990
```

Search completed: March 31, 2003, 15:36:39
Job time : 436.002 secs

```
                GenCore version 5.1.4_p5_4578
          Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 31, 2003, 13:52:37 ; Search time 196.92 Seconds
                                           (without alignments)
                                           2337.888 Million cell updates/sec

Title:          US-09-873-409-7
Perfect score:  2739
Sequence:       1 MILGILASLVNGACLPLMPL..........SESKSAVQAALEKDTPRYSF 541

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters:      1148742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115424_19290/app_query.fasta_1.7544
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09873409_@CGN_1_1_960_@runat_27032003_115424_19290
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Published_Applications_NA:*
                1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
                2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
                3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
                4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
                5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
                6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
                7:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
                8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
                9:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
                10:  /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
                11:  /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
                12:  /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
                13:  /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
                14:  /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2737	99.9	2021	10	US-09-873-409-15	Sequence 15, Appl
2	2697	98.5	3702	10	US-09-873-409-13	Sequence 13, Appl
3	2568.5	93.8	1940	10	US-09-873-409-16	Sequence 16, Appl
4	2528.5	92.3	3621	10	US-09-873-409-14	Sequence 14, Appl
5	1868	68.2	3177	10	US-09-873-409-12	Sequence 12, Appl
6	1566.5	57.2	4189	10	US-09-866-866A-5	Sequence 5, Appli
7	1555.5	56.8	3860	10	US-09-866-866A-1	Sequence 1, Appli
8	1552.5	56.7	4317	9	US-10-044-671-1	Sequence 1, Appli
9	1548	56.5	4254	10	US-09-917-800A-1424	Sequence 1424, Ap
10	1546.5	56.5	3860	10	US-09-866-866A-3	Sequence 3, Appli
11	1546.5	56.5	8630	10	US-09-306-417-1	Sequence 1, Appli
12	1546.5	56.5	8630	10	US-09-306-417-2	Sequence 2, Appli
13	1543	56.3	4425	10	US-09-769-097-3	Sequence 3, Appli
14	1541	56.3	4369	10	US-09-769-097-1	Sequence 1, Appli
15	1539	56.2	4643	9	US-10-072-621-2	Sequence 2, Appli
16	1536	56.1	4788	10	US-09-866-866A-7	Sequence 7, Appli
17	1455.5	53.1	3924	10	US-09-880-107-2299	Sequence 2299, Ap
18	1421.5	51.9	3912	10	US-09-917-800A-1560	Sequence 1560, Ap
19	1220	44.5	5010	10	US-09-917-800A-483	Sequence 483, App
20	1141.5	41.7	1175	10	US-09-873-409-11	Sequence 11, Appl
21	1101.5	40.2	2856	10	US-09-873-409-10	Sequence 10, Appl
22	1028	37.5	1810	10	US-09-749-340-4	Sequence 4, Appli
23	1028	37.5	4175	10	US-09-749-340-3	Sequence 3, Appli
24	993.5	36.3	3861	9	US-09-938-842A-2237	Sequence 2237, Ap
25	942	34.4	3861	9	US-09-938-842A-263	Sequence 263, App
26	914	33.4	4002	9	US-09-758-828-1	Sequence 1, Appli
27	914	33.4	4002	9	US-09-758-828-3	Sequence 3, Appli
28	891.5	32.5	4653	9	US-10-101-388-2	Sequence 2, Appli
29	812.5	29.7	2066	10	US-09-873-409-9	Sequence 9, Appli
30	785.5	28.7	3792	9	US-09-882-694-10	Sequence 10, Appl
31	754	27.5	2681	10	US-09-749-340-8	Sequence 8, Appli
32	754	27.5	8036	9	US-10-101-388-1	Sequence 1, Appli
33	743	27.1	3512	10	US-09-749-340-7	Sequence 7, Appli
34	738	26.9	2698	10	US-09-749-340-5	Sequence 5, Appli
35	688.5	25.1	3999	9	US-09-882-694-9	Sequence 9, Appli
36	657.5	24.0	2298	9	US-10-156-239-18	Sequence 18, Appl
37	657.5	24.0	2298	10	US-09-795-693-18	Sequence 18, Appl
38	657.5	24.0	3408	9	US-10-156-239-16	Sequence 16, Appl
39	657.5	24.0	3408	10	US-09-795-693-16	Sequence 16, Appl
40	657.5	24.0	3512	9	US-10-072-621-1	Sequence 1, Appli
41	596	21.8	2674	10	US-09-917-800A-1500	Sequence 1500, Ap
42	586.5	21.4	1749	10	US-09-815-242-9908	Sequence 9908, Ap
43	582.5	21.3	1749	10	US-09-815-242-6018	Sequence 6018, Ap
44	579.5	21.2	2247	9	US-09-870-759-40	Sequence 40, Appl
45	574.5	21.0	1737	10	US-09-815-242-8403	Sequence 8403, Ap

ALIGNMENTS

```
RESULT 1
US-09-873-409-15
; Sequence 15, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
;   LENGTH: 2021
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: No. US20020037522A1e
;   LOCATION: (723)..(723)
;   OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-15

Alignment Scores:
Pred. No.:              3.3e-288        Length:         2021
Score:                  2737.00         Matches:        541
Percent Similarity:     100.00%         Conservative:   0
```

```
Best Local Similarity:  100.00%     Mismatches:   0
Query Match:            99.93%      Indels:       0
DB:                     10          Gaps:         0

US-09-873-409-7 (1-541) x US-09-873-409-15 (1-2021)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     34 ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 93

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153

Qy     41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    154 TCTTTCTTCAGGTTGACCCTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGT 213

Qy     61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    214 TACATACAGATTTCCTTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAA 273

Qy     81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    274 CAGTTTTTTCATTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333

Qy    101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    334 GAACTTAACACTCGCATGACAGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATT 393

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAG 453

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    454 GGCTGGAAACTCACCCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA 513

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAG 633

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    634 GAGAAAGAACTTCAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTTGGTTTTATTTT 693

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    694 CCCCAGTGGCTACTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGAATCTCAAAGAT 753

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    754 GCAAAGGATTTTGGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTAC 813

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    814 TTCTTTATGAATGGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAAT 873

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    874 GGAGAACCTGGATATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGC 933

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    934 AGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCT 993

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    994 GCCTTTCATATTTTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCT 1053

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1054 GGATATAAACCTGAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTAT 1113

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1114 CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAG 1173

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1174 ACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG 1233

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1234 AGGTTATATGATCCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1293

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1294 AATGTGCGGCATTATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGG 1353

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1354 ACCACCATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1413

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1414 AGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAAT 1473

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1474 ACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATT 1533

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1534 GCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTG 1593

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAspThrProArgTyrSer 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1594 GATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTGGAGAAGGATACCCCCAGGTATTCA 1653

Qy    541 Phe 541
          |||
Db   1654 TTT 1656

RESULT 2
US-09-873-409-13
; Sequence 13, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
;   LENGTH: 3702
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: No. US20020037522A1e
;   LOCATION: (723)..(723)
;   OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-13

Alignment Scores:
Pred. No.:              1.92e-283      Length:         3702
Score:                  2697.00        Matches:        534
Percent Similarity:     100.00%        Conservative:   0
```

```
Best Local Similarity:  100.00%       Mismatches:    0
Query Match:            98.47%        Indels:        0
DB:                     10            Gaps:          0

US-09-873-409-7 (1-541) x US-09-873-409-13 (1-3702)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       34 ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 93

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153

Qy       41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      154 TCTTTCTTCAGGTTGACCCTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGT 213

Qy       61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      214 TACATACAGATTTCCTTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAA 273

Qy       81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      274 CAGTTTTTCATTCAGTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333

Qy      101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      334 GAACTTAACACTCGCATGACAGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATT 393

Qy      121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAG 453

Qy      141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      454 GGCTGGAAACTCACCCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA 513

Qy      161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573

Qy      181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAG 633

Qy      201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      634 GAGAAAGAACTTCAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTTGGTTTTATTTT 693

Qy      221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      694 CCCCAGTGGCTACTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGAATCTCAAAGAT 753

Qy      241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      754 GCAAAGGATTTGGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTAC 813

Qy      261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      814 TTCTTTATGAATGGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAAT 873

Qy      281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      874 GGAGAACCTGGATATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGC 933

Qy      301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      934 AGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCT 993

Qy      321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      994 GCCTTTCATATTTTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCT 1053

Qy      341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1054 GGATATAAACCTGAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTAT 1113

Qy      361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1114 CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAG 1173

Qy      381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1174 ACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG 1233

Qy      401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1234 AGGTTATATGATCCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1293

Qy      421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1294 AATGTGCGGCATTATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGG 1353

Qy      441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1354 ACCACCATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1413

Qy      461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1414 AGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAAT 1473

Qy      481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1474 ACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATT 1533

Qy      501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1534 GCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTG 1593

Qy      521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
            ||||||||||||||||||||||||||||||||||||||||||
Db     1594 GATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTGGAGAAG 1635

RESULT 3
US-09-873-409-16
; Sequence 16, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-409-16

Alignment Scores:
Pred. No.:              6.98e-270     Length:        1940
Score:                  2568.50       Matches:       514
Percent Similarity:     95.01%        Conservative:  0
Best Local Similarity:  95.01%        Mismatches:    0
Query Match:            93.78%        Indels:        27
DB:                     10            Gaps:          1

US-09-873-409-7 (1-541) x US-09-873-409-16 (1-1940)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      34 ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 93

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153

Qy      41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     154 TCTTTCTTCAGGTTGACCCTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGT 213

Qy      61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     214 TACATACAGATTTCCTTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAA 273

Qy      81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     274 CAGTTTTTTCATTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333

Qy     101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     334 GAACTTAACACTCGCATGACAGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATT 393

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAG 453

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     454 GGCTGGAAACTCACCCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA 513

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAG 633

Qy     201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
           |||||||||||||||
Db     634 GAGAAAGAACTTCAA--------------------------------------------- 648

Qy     221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                                    ||||||||||||||||||||||||
Db     649 ------------------------------------AGGTATACACAGAATCTCAAAGAT 672

Qy     241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     673 GCAAAGGATTTTGGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTAC 732

Qy     261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     733 TTCTTTATGAATGGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAAT 792

Qy     281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     793 GGAGAACCTGGATATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGC 852

Qy     301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     853 AGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCT 912

Qy     321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     913 GCCTTTCATATTTTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCT 972

Qy     341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     973 GGATATAAACCTGAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTAT 1032

Qy     361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1033 CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAG 1092

Qy     381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1093 ACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG 1152

Qy     401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1153 AGGTTATATGATCCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1212

Qy     421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1213 AATGTGCGGCATTATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGG 1272

Qy     441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1273 ACCACCATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1332

Qy     461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1333 AGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAAT 1392

Qy     481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1393 ACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATT 1452

Qy     501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1453 GCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTG 1512

Qy     521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAspThrProArgTyrSer 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1513 GATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTGGAGAAGGATACCCCCAGGTATTCA 1572

Qy     541 Phe 541
           |||
Db    1573 TTT 1575
```

RESULT 4
US-09-873-409-14
; Sequence 14, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
;   LENGTH: 3621
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-409-14

Alignment Scores:

Pred. No.:	4.18e-265	Length:	3621
Score:	2528.50	Matches:	507
Percent Similarity:	94.94%	Conservative:	0
Best Local Similarity:	94.94%	Mismatches:	0
Query Match:	92.31%	Indels:	27
DB:	10	Gaps:	1

US-09-873-409-7 (1-541) x US-09-873-409-14 (1-3621)

```
Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      34 ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 93

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153

Qy      41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     154 TCTTTCTTCAGGTTGACCCTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGT 213

Qy      61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     214 TACATACAGATTTCCTTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAA 273

Qy      81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     274 CAGTTTTTCATTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333

Qy     101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     334 GAACTTAACACTCGCATGACAGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATT 393

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAG 453

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     454 GGCTGGAAACTCACCCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA 513

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAG 633

Qy     201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
           ||||||||||||||
Db     634 GAGAAAGAACTTCAA--------------------------------------------- 648

Qy     221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                                ||||||||||||||||||||||||
Db     649 ------------------------------------AGGTATACACAGAATCTCAAAGAT 672

Qy     241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     673 GCAAAGGATTTTGGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTAC 732

Qy     261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     733 TTCTTTATGAATGGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAAT 792

Qy     281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     793 GGAGAACCTGGATATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGC 852

Qy     301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     853 AGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCT 912

Qy     321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     913 GCCTTTCATATTTTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCT 972

Qy     341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     973 GGATATAAACCTGAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTAT 1032

Qy     361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1033 CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAG 1092

Qy     381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1093 ACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG 1152

Qy     401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1153 AGGTTATATGATCCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1212

Qy     421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1213 AATGTGCGGCATTATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGG 1272

Qy     441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1273 ACCACCATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1332

Qy     461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1333 AGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAAT 1392

Qy     481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1393 ACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATT 1452

Qy     501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1453 GCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTG 1512

Qy     521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
           ||||||||||||||||||||||||||||||||||||||||||
Db    1513 GATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTGGAGAAG 1554
```

RESULT 5
US-09-873-409-12
; Sequence 12, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
; LENGTH: 3177
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: No. US20020037522A1e
; LOCATION: (198)..(198)
; OTHER INFORMATION: n at position 198 represents any nucleotide (A, T, C or G)
US-09-873-409-12

Alignment Scores:

Pred. No.:	2.97e-193	Length:	3177
Score:	1868.00	Matches:	370
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	68.20%	Indels:	0
DB:	10	Gaps:	0

US-09-873-409-7 (1-541) x US-09-873-409-12 (1-3177)

```
Qy     165 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 184
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCA 60

Qy     185 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 204
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAGGAGAAAGAACTT 120

Qy     205 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 224
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    121 CAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTTGGTTTATTTCCCCAGTGGCTA 180

Qy    225 LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 244
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 CTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGAATCTCAAAGATGCAAAGGATTTT 240

Qy    245 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 264
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 GGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAAT 300

Qy    265 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 284
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 GGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGA 360

Qy    285 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 304
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 TATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATT 420

Qy    305 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 324
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATT 480

Qy    325 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 344
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 TTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCT 540

Qy    345 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 364
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    541 GAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCA 600

Qy    365 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 384
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 TCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTG 660

Qy    385 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 404
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    661 GTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 720

Qy    405 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 424
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    721 CCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCAT 780

Qy    425 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 444
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    781 TATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGT 840

Qy    445 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 464
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    841 AACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGG 900

Qy    465 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 484
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    901 GAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGG 960

Qy    485 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 504
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTA 1020

Qy    505 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 524
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 GTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGC 1080

Qy    525 LysSerAlaValGlnAlaAlaLeuGluLys 534
          ||||||||||||||||||||||||||||||
Db   1081 AAGTCAGCTGTTCAAGCTGCACTGGAGAAG 1110
```

RESULT 6
US-09-866-866A-5
; Sequence 5, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 4189
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-866-866A-5

```
Alignment Scores:
Pred. No.:              3.15e-160     Length:         4189
Score:                  1566.50       Matches:        307
Percent Similarity:     72.63%        Conservative:   99
Best Local Similarity:  54.92%        Mismatches:     94
Query Match:            57.19%        Indels:         59
DB:                     10            Gaps:           6

US-09-873-409-7 (1-541) x US-09-866-866A-5 (1-4189)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          |||||||||||  |||||:::::::::::|||      ||||||||:::   |||
Db    148 ATGATTCTGGGAACTCTCGCTGCTATTATCCATGGAACATTACTTCCCCTCTTGATGCTG 207

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
          |||   |||   |||:::|||:::
Db    208 GTGTTTGGAAACATGACAGATAGTTTACAAAAGCAGAAGCCAGTATTCTGCCAAGCATT 267

Qy     31 -----------------SerGlyCysLeuValGlnThrAsnThr--------------- 39
                           :::   |||:::  :::|||:::
Db    268 ACTAATCAAAGTGGACCCAACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAAGAG 327

Qy     40 ---------TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56
                   |||::::::               ||||||  ||||||||
Db    328 ATGGCCATATACGCCTAC---------------TATTACACCGGGATTGGTGCTGGTGTG 372

Qy     57 LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys 76
          ||||||   ||||||||:::|||||||||     :::   |||   ||||||
Db    373 CTCATAGTTGCCTACATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGAAGACAGATACAC 432

Qy     77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96
          :::||||||::::::|||||||||:::::::::   |||::::||||||||||||||||
Db    433 AAGATTAGGCAGAAGTTTTTCCATGCTATAATGAATCAGGAGATAGGCTGGTTTGATGTG 492

Qy     97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
             |||:::||||||||||||||||||:::|||   |||:::   ||||||:::||||||
Db    493 CATGATGTTGGGGAGCTCAACACCCGGCTCACAGATGATGTCTCCAAAATTAATGACGGA 552

Qy    116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
          |||||||||||||||   :::   ||||||::::::::::||||||      |||
Db    553 ATTGGTGACAAAATTGGGATGTTTTTTCAGTCCATAACCACATTTTTAGCCGGTTTTATC 612

Qy    136 ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
          :::|||   :::   ||||||||||||||||||||||||   |||:::   ||||||||||||
Db    613 ATAGGATTTATAAGTGGTTGGAAGCTAACCCTTGTCATTTTGGCTGTCAGCCCTCTTATT 672

Qy    156 MetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
                |||:::|||      ::::::::::::   |||   |||:::|||||||||
Db    673 GGATTGTCATCTGCTTTGTGGGCAAAGGTATTGACTTCATTTACTAATAAGGAACTCCAG 732

Qy    176 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIle 195
          ||||||:::|||||||||||||||||||||||||||||||::::::||||||||||||||
Db    733 GCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT 792

Qy    196 AlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyr 215
```

```
              ||||||       |||:::|||||||||:::
Db      793 GCCTTTGGAGGACAACAGAAGGAACTTGAA----------------------------- 822

Qy      216 AlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThr 235
                                                                |||||| 
Db      823 -----------------------------------------------------AGGTACAAT 831

Qy      236 GlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSer 255
            :::|||||::::::||||:::     |||||||||:::   |||   :::   :::|||
Db      832 AAAAATTTAGAAGAAGCTAAAAATGTTGGCATAAAGAAAGCTATCACAGCCAGCATTTCG 891

Qy      256 LeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThr 275
            :::|||     |||      :::      :::|||   |||||||||||||||||||
Db      892 ATAGGCATTGCCTACCTGTTGGTCTATGCATCATATGCACTGGCATTCTGGTATGGACA 951

Qy      276 SerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePhe 295
            ||||||:::|||:::   |||       |||:::||||||   ||||||   |||||||||
Db      952 TCCTTGGTCCTCTCAAATGAA------TATTCTATTGGAGAAGTGCTTACTGTCTTCTTC 1005

Qy      296 SerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAla 315
            |||::::::         ::::::   ||||||        |||:::   |||   ||||||
Db     1006 TCTATTTTGTTGGGGACTTTTAGTATTGGACACTTGGCCCCAAACATAGAAGCCTTTGCA 1065

Qy      316 IleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAsp 335
               |||||||||||||||||   ||||||:::::||||||   :::|||||||||||
Db     1066 AACGCACGAGGGGCAGCCTTTGAAATCTTCAAGATAATTGATAACGAGCCAAGCATTGAC 1125

Qy      336 AsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsn 355
            :::|||||||||   ||||||||||||:::||||||   |||      :::||||||||||||
Db     1126 AGCTTCTCAACAAAGGGCTACAAACCAGACAGTATAATGGGAAACTTAGAGTTTAAAAAT 1185

Qy      356 ValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArg 375
            |||   ||||||||||||||||||          ::::::||||||||||||||||||:::
Db     1186 GTTCACTTCAACTACCCATCGAGAAGCGAAGTTCAGATCTTGAAGGGCCTCAATCTGAAG 1245

Qy      376 IleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVal 395
            :::|||||||||:::||||||||||||||||||      :::|||   ||||||||||||
Db     1246 GTGAAGAGCGGACAGACGGTGGCCTTGGTTGGCAACAGTGGCTGTGGAAAAAGCACAACT 1305

Qy      396 ValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsn 415
            ||||||||||:::||||||||||||||||||   :::|||   :::   :::|||
Db     1306 GTCCAGCTGATGCAGAGGCTCTACGACCCCCTGGAGGGCGTGGTCAGTATCGACGGACAA 1365

Qy      416 AspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGlu 435
            |||||||||   :::|||||||||:::   |||:::   ||||||||||||||||||||
Db     1366 GACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCATTGGTGTGGTGAGTCAGGAA 1425

Qy      436 ProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThr 455
            ||||||||||||   ||||||||||:::   ||||||:::||||||||:::|||||||||
Db     1426 CCTGTGCTGTTTGCCACCACGATCGCCGAGAACATTCGCTATGGCCGAGAAGATGTCACC 1485

Qy      456 AspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhe 475
               :::|||:::|||:::|||   :::||||||||||||||||||||||||||:::
Db     1486 ATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTCATCATGAAACTG 1545

Qy      476 ProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLys 495
            |||::::::|||:::||||||||||||||||||:::||||||:::||||||||||||||
Db     1546 CCCCACCAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAA 1605

Qy      496 GlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGlu 515
            |||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db     1606 CAGAGAATCGCCATTGCCCGGGCCCTGGTCCGCAATCCCAAGATCCTTTTGTTGGACGAG 1665

Qy      516 AlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
            ||||||||||||||||||:::||||||::::::   |||||||||||||||:::|||
Db     1666 GCCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTGGTGCAGGCCGCACTGGATAAG 1722
```

RESULT 7
US-09-866-866A-1
; Sequence 1, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
;   LENGTH: 3860
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-866-866A-1

Alignment Scores:			
Pred. No.:	4.38e-159	Length:	3860
Score:	1555.50	Matches:	306
Percent Similarity:	73.29%	Conservative:	100
Best Local Similarity:	55.23%	Mismatches:	99
Query Match:	56.79%	Indels:	49
DB:	10	Gaps:	6

US-09-873-409-7 (1-541) x US-09-866-866A-1 (1-3860)

```
Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            |||::::::|||   ||||||::::::::::::||||||   ||||||||||||   |||
Db      151 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210

Qy       21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
            |||   |||||||||:::|||                          :::|||:::|||
Db      211 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 270

Qy       32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe----------------- 43
                     ::::::        |||      ||||||
Db      271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330

Qy       44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                  |||         ||||||   |||||||||         |||:::      |||
Db      331 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 390

Qy       62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
            ||||||:::|||   |||   :::   |||      ||||||      :::|||||||||
Db      391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450

Qy       82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
            ||||||||:::::::::   |||:::||||||||||||||      |||:::||||||
Db      451 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 510

Qy      102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
            ||||||||||||:::|||   |||:::   ||||||::::::||||||||||||||||||
Db      511 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGGTATTGGTGACAAAATT 570

Qy      121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
               :::   ||||||:::|||:::||||||      |||      ||||||      :::
Db      571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 630

Qy      141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
            ||||||||||||||||||   |||:::   ||||||::::::         |||||||||
Db      631 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

Qy      161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
               ::::::::::::   |||   |||   |||||||||      ||||||:::||||||
Db      691 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy      181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
            ||||||||||||||||||||||||::::::||||||||||||||||||||      |||
Db      751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810
```

```
Qy     201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
           :::||||||||:::
Db     811 AAGAAAGAACTTGAA--------------------------------------------- 825

Qy     221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                           ||||||  :::||||||::::::
Db     826 -----------------------------------AGGTACAACAAAAATTTAGAAGAA 849

Qy     241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
           ||||||    |||||||||:::  |||   :::   :::|||:::||||||    :::
Db     850 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 909

Qy     261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
                 :::    :::|||  ||||||||||||||||||||:::|||:::|||:::
Db     910 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 969

Qy     281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
           ||||||     |||:::||||||   ||||||  ||||||||||||||||:::
Db     970 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1023

Qy     301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
           ::::::   :::|||  |||   |||      |||  ||||||   ||||||||||||
Db    1024 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1083

Qy     321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
           |||:::   ||||||::::::||||||   ||||||||||||||::::::|||    :::
Db    1084 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1143

Qy     341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
           |||:::||||||::::::|||:::|||    :::||||||:::||||||   |||:::|||
Db    1144 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1203

Qy     361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
           |||||||||      :::||||||||||||||||||||||||:::::::::||||||:::
Db    1204 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1263

Qy     381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
           ||||||||||||||||||   :::|||   |||||||||||    ||||||||||:::|||
Db    1264 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1323

Qy     401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
           |||||||||||||||   :::|||   :::   ||||||      ||||||||   :::
Db    1324 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1383

Qy     421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
           |||||||||      |||:::   |||||||||||||||||||||||||||||||
Db    1384 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1443

Qy     441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
           |||||||||:::     ||||||:::||||||||||::::::||||||    :::|||:::|||
Db    1444 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1503

Qy     461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
           :::|||    :::|||||||||||||||||||||||||||:::   |||:::||||||:::
Db    1504 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1563

Qy     481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
           ||||||||||||||||:::||||||||||:::|||||||||||||||||||||||||||||
Db    1564 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 1623

Qy     501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
           |||||||||||||||||||||||||||||||:::||||||||||||||||||||||||||
Db    1624 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 1683

Qy     521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
           |||:::||||||::::::   ||||||   ||||||:::|||
Db    1684 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 1725
```

RESULT 8
US-10-044-671-1
; Sequence 1, Application US/10044671
; Patent No. US20020177147A1
; GENERAL INFORMATION:
; APPLICANT: Washington State University Research Foundation
; APPLICANT: Mealey, Katrina
; APPLICANT: Bentjen, Steven
; TITLE OF INVENTION: MDR1 VARIANTS AND METHODS FOR THEIR USE
; FILE REFERENCE: 4630-61733
; CURRENT APPLICATION NUMBER: US/10/044,671
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/261,578
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/314,829
; PRIOR FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 4317
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (70)..(3912)
; OTHER INFORMATION:
US-10-044-671-1

Alignment Scores:

Pred. No.:	1.1e-158	Length:	4317
Score:	1552.50	Matches:	302
Percent Similarity:	71.94%	Conservative:	98
Best Local Similarity:	54.32%	Mismatches:	105
Query Match:	56.68%	Indels:	51
DB:	9	Gaps:	5

US-09-873-409-7 (1-541) x US-10-044-671-1 (1-4317)

```
Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
              |||::::::|||    :::|||::::::::::::||||||  ||||||||||||   |||
Db     217 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 276

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
           |||   |||   |||:::|||:::   :::      :::
Db     277 GTTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 336

Qy      35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                                :::      ||||||   |||
Db     337 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 396

Qy      45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                          ||||||   |||||||||            |||:::
Db     397 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 456

Qy      60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
              |||||||||:::|||   |||    :::   |||    ||||||      :::||||||
Db     457 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 516

Qy      80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
           ||||||||||||||:::::::::   |||:::||||||||||||||||      |||:::
Db     517 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 576

Qy     100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
           ||||||||||||||||||:::|||   |||:::   ||||||:::::::||||||||||||
Db     577 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 636

Qy     119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
           |||:::   :::   ||||||:::::::::||||||      |||      ||||||
Db     637 AAAGTTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 696

Qy     139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
              ||||||||||||||||||||||||   |||:::   ||||||:::::::      |||
Db     697 ACACCTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 756

Qy     159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
           ||||||      ::::::::::::   |||   |||   |||||||||   ||||||:::
Db     757 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 816
```

```
Qy     179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
           ||||||||||||||||||||||||||||||||:::::|||||||||||||||||||||||
Db     817 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 876

Qy     199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
              |||:::||||||||:::
Db     877 GGACAAAAGAAAGAACTTGAA--------------------------------------- 897

Qy     219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
                                                  ||||||   :::||||||
Db     898 ---------------------------------------AGGTACAACAAAAATTTA 915

Qy     239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
           ::::::||||||      |||||||||:::   |||   :::   :::|||:::||||||
Db     916 GAAGAAGCTAAAAGAATTGGGATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCC 975

Qy     259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
              :::      :::      :::|||   |||||||||||||||||||||||||||:::
Db     976 GCTTTCTTATTGATCTATGCATCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTC 1035

Qy     279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
           |||:::   |||      |||||||||||   ||||||   ||||||||||||||:::
Db    1036 CTCTCCAGTGAA------TATACTATTGGACAGGTACTCACTGTCTTCTTTTCTGTATTA 1089

Qy     299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
                 ::::::   ||||||   |||   |||      |||   ||||||   ||||||
Db    1090 ATTGGGGCTTTTAGTATTGGACAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGA 1149

Qy     319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
           |||||||||:::   ||||||:::::||||||   |||||||||||||||::::::|||
Db    1150 GGAGCAGCTTATGAAATCTTCAAGATAATTGACAATAAACCAAGCATTGACAGCTATTCG 1209

Qy     339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
              :::|||:::||||||::::::|||:::|||   :::||||||||||||||   |||
Db    1210 AAGAGTGGACATAAACCAGATAATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTC 1269

Qy     359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
           :::||||||||||||   :::||||||||||||||||||||||||:::::::::|||
Db    1270 AGTTACCCTTCTCGAAAAGAAGTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGT 1329

Qy     379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
           |||:::|||||||||||||||||||   :::|||   ||||||||||||   |||||||||
Db    1330 GGGCAGACAGTGGCGCTGGTTGGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTG 1389

Qy     399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
           :::||||||||||||||||||   ||||||   :::   :::|||      |||||||||
Db    1390 ATGCAGAGGCTCTATGACCCCACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGG 1449

Qy     419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
              :::||||||||||||   |||:::         ||||||||||||||||||||||||
Db    1450 ACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTG 1509

Qy     439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
           |||   ||||||||:::   |||||||:::|||||||||||:::::||||||   :::|||
Db    1510 TTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAG 1569

Qy     459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
           :::|||:::|||   :::||||||||||||||||||||||||:::      |||||||||
Db    1570 ATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTTATCATGAAACTACCTAATAAA 1629

Qy     479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
           |||:::|||||||||||||||:::|||||||:::::::|||||||||||||||||||||||
Db    1630 TTTGACACTCTGGTTGGAGAGAGAGGGGCCCGGCTGAGTGGTGGACAGAAACAGAGAATC 1689

Qy     499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
           |||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db    1690 GCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCA 1749

Qy     519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
           |||||||||:::||||||::::::   ||||||   ||||||:::|||
Db    1750 GCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCCCTGGATAAG 1797
```

RESULT 9

```
US-09-917-800A-1424
; Sequence 1424, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
;  APPLICANT: Mendrick, Donna
;  APPLICANT:  Porter, Mark
;  APPLICANT:  Johnson, Kory
;  APPLICANT:  Castle, Arthur
;  APPLICANT:  Elashoff, Michael
;  APPLICANT:  Gene Logic, Inc.
;  TITLE OF INVENTION: Molecular Toxicology Modeling
;  FILE REFERENCE: 44921-5038-US
;  CURRENT APPLICATION NUMBER: US/09/917,800A
;  CURRENT FILING DATE:  2001-07-31
;  PRIOR APPLICATION NUMBER: US 60/222,040
;  PRIOR FILING DATE: 2000-07-31
;  PRIOR APPLICATION NUMBER: US 60/222,880
;  PRIOR FILING DATE: 2000-11-02
;  PRIOR APPLICATION NUMBER: US 60/290,029
;  PRIOR FILING DATE: 2001-05-11
;  PRIOR APPLICATION NUMBER: US 60/290,645
;  PRIOR FILING DATE: 2001-05-15
;  PRIOR APPLICATION NUMBER: US 60/292,336
;  PRIOR FILING DATE: 2001-05-22
;  PRIOR APPLICATION NUMBER: US 60/295,798
;  PRIOR FILING DATE: 2001-06-06
;  PRIOR APPLICATION NUMBER: US 60/297,457
;  PRIOR FILING DATE: 2001-06-13
;  PRIOR APPLICATION NUMBER: US 60/298,884
;  PRIOR FILING DATE: 2001-06-19
;  PRIOR APPLICATION NUMBER: US 60/303,459
;  PRIOR FILING DATE: 2001-07-09
;  NUMBER OF SEQ ID NOS: 1740
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1424
;   LENGTH: 4254
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020119462A1 M81855
US-09-917-800A-1424

Alignment Scores:
Pred. No.:              3.34e-158      Length:        4254
Score:                  1548.00        Matches:       307
Percent Similarity:     72.25%         Conservative:  94
Best Local Similarity:  55.32%         Mismatches:    104
Query Match:            56.52%         Indels:        50
DB:                     10             Gaps:          5

US-09-873-409-7 (1-541) x US-09-917-800A-1424 (1-4254)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||   ||||||   ||||||::::::::::::|||   |||||||||:::   |||
Db     247 ATGGCTCTGGGAACTCTCGCTGCTATCATCCACGGAACCCTGCTTCCCCTCCTGATGCTG 306

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
           |||   |||   |||:::|||:::
Db     307 GTGTTCGGATACATGACAGATAGTTTACCCCAAGCAGAGACCCGCATTCTGACCGAGCG 366

Qy      31 ---------SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu------ 45
                          :::      |||:::|||   |||         |||
Db     367 ATTACTAATCAAAGTGAAATCAACAGTACACATACCGTCAGCGACACGAGTCTGGAGGAG 426

Qy      46 ---------------ThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
                          ||||||   |||||||||         ||||||
Db     427 GACATGGCCATGTATGCCTACTATTACACGGGCATTGGTGCCGGTGTGCTCATCGTTGCC 486

Qy      61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
           |||||||:::||||||||   :::   |||   ||||||         :::||||||:::
Db     487 TACATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAG 546

Qy      81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
```

```
Db    547 AAGTTTTTCCATGCCATCATGAATCAGGAGATAGGCTGGTTTGACGTGAATGACGCTGGG 606

Qy    101 GluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLys 119
Db    607 GAGCTCAACACCCGTCTCACAGATGACGTCTCCAAAATTAATGACGGAATTGGTGACAAA 666

Qy    120 IleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuVal 139
Db    667 CTTGGAATGTTCTTTCAGTCCATAACGACATTTTCAGCCGGTTTTATAATAGGATTTATA 726

Qy    140 LysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAla 159
Db    727 AGTGGTTGGAAGCTAACCCTTGTAATTTTGGCCGTCAGCCCTCTTATTGGGTTGTCATCT 786

Qy    160 AlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLys 179
Db    787 GCCATGTGGGCAAAGGTACTGACTTCATTTACTAATAAGGAACTCCAGGCTTATGCGAAA 846

Qy    180 AlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAla 199
Db    847 GCTGGAGCAGTTGCCGAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCGTTTGGAGGA 906

Qy    200 GlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyr 219
Db    907 CAAAAGAAGGAACTTGAA------------------------------------------ 924

Qy    220 PheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLys 239
Db    925 ---------------------------------------AGGTACAATAAAAATTTAGAA 945

Qy    240 AspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaVal 259
Db    946 GAAGCTAAAAGAGTTGGCATAAAGAAAGCCATCACGGCCAACATTTCCATAGGTATTGCC 1005

Qy    260 TyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeu 279
Db   1006 TACCTGTTGGTCTATGCGTCTTATGCACTGGCATTCTGGTATGGGACCTCCTTGGTCCTC 1065

Qy    280 AsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHis 299
Db   1066 TCAAATGAA------TATTCTATTGGACAAGTGCTTACCGTCTTCTTCTCTATTTTATTG 1119

Qy    300 SerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGly 319
Db   1120 GGACTTTCAGTATTGGACATTTAGCCCCAAACATAGAAGCCTTTGCAAATGCAAGAGGG 1179

Qy    320 AlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThr 339
Db   1180 GCAGCCTATGAAATCTTCAAGATAATTGATAATGAGCCAAGCATCGACAGCTTCTCAACC 1239

Qy    340 AlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsn 359
Db   1240 AAGGGACACAAACCAGACAGTATAATGGGAAATTTGGAATTTAAAAATGTTTACTTCAAC 1299

Qy    360 TyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGly 379
Db   1300 TACCCATCACGAAGTGAAGTTAAGATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGG 1359

Qy    380 GluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeu 399
Db   1360 CAGACGGTAGCCCTGGTTGGCAACAGTGGCTGTGGAAAAGCACAACTGTCCAGCTGCTG 1419

Qy    400 GlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAla 419
Db   1420 CAGAGGCTCTACGACCCCATAGAGGGCGAGGTCAGTATTGACGGACAGGACATCAGGACC 1479

Qy    420 LeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPhe 439
Db   1480 ATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTT 1539

Qy    440 GlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMet 459
Db   1540 GCCACCACGATTGCCGAAAACATTCGCTATGGCCGAGAAAACGTCACCATGGATGAGATA 1599

Qy    460 GluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPhe 479
Db   1600 GAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTCATCATGAAACTGCCCCACAAATTT 1659

Qy    480 AsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAla 499
Db   1660 GACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCC 1719

Qy    500 IleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAla 519
Db   1720 ATTGCCCGGGCCCTGGTCCGCAACCCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCC 1779

Qy    520 LeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db   1780 TTGGACACAGAAAGCGAAGCCGTGGTTCAGGCCGCTCTGGATAAG 1824
```

RESULT 10
US-09-866-866A-3
; Sequence 3, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 3860
; TYPE: DNA
; ORGANISM: homo sapiens
US-09-866-866A-3

Alignment Scores:

Pred. No.:	4.18e-158	Length:	3860
Score:	1546.50	Matches:	305
Percent Similarity:	73.10%	Conservative:	100
Best Local Similarity:	55.05%	Mismatches:	100
Query Match:	56.46%	Indels:	49
DB:	10	Gaps:	6

US-09-873-409-7 (1-541) x US-09-866-866A-3 (1-3860)

```
Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    151 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210

Qy     21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
Db    211 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 270

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db    271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    331 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 390

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450
```

```
Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    451 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 510

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    511 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 570

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 630

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    631 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    691 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db    811 AAGAAAGAACTTGAA--------------------------------------------- 825

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db    826 ------------------------------------AGGTACAACAAAAATTTAGAAGAA 849

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db    850 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 909

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db    910 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 969

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db    970 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1023

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db   1024 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1083

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   1084 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1143

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   1144 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1203

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   1204 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1263

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   1264 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1323

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   1324 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1383

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   1384 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1443

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   1444 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1503

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   1504 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1563

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   1564 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 1623

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   1624 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 1683

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db   1684 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 1725
```

RESULT 11
US-09-306-417-1
; Sequence 1, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 8630
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: proviral
;   OTHER INFORMATION: plasmid DNA
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(160)
;   OTHER INFORMATION: plasmid backbone (pUC)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (161)..(677)
;   OTHER INFORMATION: 5'-LTR
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (532)..(1219)
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: (1220)..(5062)
;   OTHER INFORMATION: m4 mdr-1 cDNA
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5215)..(5774)
;   OTHER INFORMATION: 3'-LTR
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5775)..(8630)
;   OTHER INFORMATION: plasmid backbone (pUC)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(8630)
;   OTHER INFORMATION: retroviral expression vector SFbeta71m4
US-09-306-417-1

```
Alignment Scores:
Pred. No.:              1.46e-157    Length:        8630
Score:                  1546.50      Matches:       305
Percent Similarity:     73.10%       Conservative:  100
```

```
Best Local Similarity:  55.05%     Mismatches:    100
Query Match:            56.46%     Indels:         49
DB:                     10         Gaps:            6

US-09-873-409-7 (1-541) x US-09-306-417-1 (1-8630)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    1370 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 1429

Qy      21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db    1430 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 1489

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db    1490 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 1549

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    1550 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 1609

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    1610 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 1669

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    1670 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 1729

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    1730 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 1789

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    1790 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1849

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    1850 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1909

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    1910 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1969

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    1970 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 2029

Qy     201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db    2030 AAGAAAGAACTTGAA--------------------------------------------- 2044

Qy     221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db    2045 -----------------------------------AGGTACAACAAAAATTTAGAAGAA 2068

Qy     241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db    2069 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 2128

Qy     261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db    2129 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 2188

Qy     281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db    2189 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 2242

Qy     301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db    2243 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 2302

Qy     321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db    2303 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 2362

Qy     341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db    2363 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 2422

Qy     361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db    2423 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 2482

Qy     381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db    2483 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 2542

Qy     401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db    2543 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 2602

Qy     421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db    2603 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 2662

Qy     441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db    2663 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 2722

Qy     461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db    2723 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 2782

Qy     481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db    2783 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 2842

Qy     501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db    2843 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 2902

Qy     521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db    2903 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 2944
```

RESULT 12
US-09-306-417-2
; Sequence 2, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 8630
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: proviral
;   OTHER INFORMATION: plasmid DNA
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(8630)
;   OTHER INFORMATION: retroviral expression vector SFbeta91mSA1
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(160)

```
;    OTHER INFORMATION: plasmid backbone (pUC)
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (161)..(677)
;    OTHER INFORMATION: 5'-LTR
;    FEATURE:
;    NAME/KEY: 5'UTR
;    LOCATION: (532)..(1219)
;    FEATURE:
;    NAME/KEY: mat_peptide
;    LOCATION: (1220)..(5062)
;    OTHER INFORMATION: mSA1 mdr1 cDNA
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (5215)..(5774)
;    OTHER INFORMATION: 3'-LTR
;    FEATURE:
;    NAME/KEY: misc_feature
;    LOCATION: (5775)..(8630)
;    OTHER INFORMATION: plasmid backbone (pUC)
US-09-306-417-2

Alignment Scores:
Pred. No.:              1.46e-157     Length:        8630
Score:                  1546.50       Matches:       305
Percent Similarity:     73.10%        Conservative:  100
Best Local Similarity:  55.05%        Mismatches:    100
Query Match:            56.46%        Indels:        49
DB:                     10            Gaps:          6

US-09-873-409-7 (1-541) x US-09-306-417-2 (1-8630)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db   1370 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 1429

Qy     21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db   1430 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 1489

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db   1490 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 1549

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db   1550 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 1609

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db   1610 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 1669

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db   1670 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 1729

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db   1730 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 1789

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db   1790 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1849

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db   1850 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1909

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db   1910 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1969

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db   1970 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 2029

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db   2030 AAGAAAGAACTTGAA--------------------------------------------- 2044

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db   2045 -----------------------------------AGGTACAACAAAAATTTAGAAGAA 2068

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db   2069 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 2128

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db   2129 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 2188

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db   2189 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 2242

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db   2243 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 2302

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   2303 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 2362

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   2363 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 2422

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   2423 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 2482

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   2483 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 2542

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   2543 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 2602

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   2603 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 2662

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   2663 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 2722

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   2723 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 2782

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   2783 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 2842

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   2843 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 2902

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db   2903 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 2944

RESULT 13
US-09-769-097-3
```

```
; Sequence 3, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
;  APPLICANT: Kimberly Anne Brun
;  APPLICANT:  Richard James Chenery
;  APPLICANT:  Harma Ellens
;  APPLICANT:  John Anthony Feild
;  APPLICANT:  Lin Yue
;  TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
;  TITLE OF INVENTION:  ENCODING RAT MDR1A AND SCREENING METHODS THEREOF
;  FILE REFERENCE: GP-50009-C2
;  CURRENT APPLICATION NUMBER: US/09/769,097
;  CURRENT FILING DATE:  2001-01-25
;  PRIOR APPLICATION NUMBER: 09/208,809
;  PRIOR FILING DATE: 1998-12-09
;  PRIOR APPLICATION NUMBER: 09/156,800
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: US99/20770
;  PRIOR FILING DATE: 1999-09-10
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 4425
;   TYPE: DNA
;   ORGANISM: RATTUS RATTUS
US-09-769-097-3

Alignment Scores:
Pred. No.:              1.24e-157     Length:         4425
Score:                  1543.00       Matches:        300
Percent Similarity:     73.04%        Conservative:   101
Best Local Similarity:  54.64%        Mismatches:     102
Query Match:            56.33%        Indels:         46
DB:                     10            Gaps:           5

US-09-873-409-7 (1-541) x US-09-769-097-3 (1-4425)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    499 ATGCTGCTGGGAACTCTGGCGGCCATTATCCATGGAATTGCGCTCCCACTTATGATGCTG 558

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db    559 GTCTTTGGAGACATGACAGATAGCTTTGCAAAT------GTAGGAAACAACCGTAGTATG 612

Qy     41 SerPhePheArgLeuThr------------------------------------------ 46
Db    613 AGTTTCTACAATGCTACAGACATATATGCCAAGCTGGAGGACGAAATGACCACGTACGCC 672

Qy     47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
Db    673 TACTATTACACGGGCATTGGTGCCGGTGTGCTCATCGTTGCCTACATCCAGGTTTCCACT 732

Qy     67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
Db    733 TGGTGCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAAGTTTTTCCATGCCATC 792

Qy     87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
Db    793 ATGAATCAGGAGATAGGCTGGTTTGACGTGCATGACGTTGGGGAGCTCAACACCCGGCTC 852

Qy    107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
Db    853 ACAGATGACGTCTCCAAAATTAATGAAGGAATTGGTGACAAAATTGGAATGTTCTTTCAG 912

Qy    126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
Db    913 GCAATGGCAACATTTTTTGGTGGTTTTATAATAGGATTTACTCGCGGCTGGAAGCTAACT 972

Qy    146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
Db    973 CTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTGTCAGCTGGTATTTGGGCAAAGATA 1032

Qy    166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
Db   1033 TTGTCTTCATTTACTGATAAGGAACTCCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAA 1092

Qy    186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
Db   1093 GAAGTCTTAGCAGCCATCAGAACTGTGATTGCCTTTGGAGGACAAAAGAAGGAACTTGAA 1152

Qy    206 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeu 225
Db   1152 ------------------------------------------------------------ 1152

Qy    226 SerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 245
Db   1153 ---------------------AGGTACAATAACAATTTGGAAGAAGCTAAAAGGCTTGGG 1191

Qy    246 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 265
Db   1192 ATAAAGAAAGCTATCACGGCCAACATTTCCATGGGTGCAGCTTTTCTGCTTATCTATGCA 1251

Qy    266 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 285
Db   1252 TCATATGCTCTGGCATTCTGGTATGGGACTTCCTTGGTCATCTCAAAAGAA------TAC 1305

Qy    286 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 305
Db   1306 ACTATTGGACAAGTGCTCACTGTCTTTTTTTCTGTATTAATTGGAGCATTCAGTGTTGGG 1365

Qy    306 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 325
Db   1366 CAGGCATCTCCAAATATTGAAGCCTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTC 1425

Qy    326 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 345
Db   1426 AGTATAATTGATAATAAGCCCAGTATAGACAGCTTCTCAAAGAGTGGGCACAAACCCGAC 1485

Qy    346 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 365
Db   1486 AACATACAAGGAAATTTGGAATTCAAAAATATTCACTTCAGTTACCCGTCTCGAAAAGAC 1545

Qy    366 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 385
Db   1546 GTTCAGATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTGGTT 1605

Qy    386 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 405
Db   1606 GGCAACAGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCC 1665

Qy    406 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 425
Db   1666 ATAGAGGGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTG 1725

Qy    426 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 445
Db   1726 CGGGAAATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCCACCACAATTGCCGAA 1785

Qy    446 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 465
Db   1786 AACATTCGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAA 1845

Qy    466 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 485
Db   1846 GCCAATGCCTATGATTTCATCATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAG 1905

Qy    486 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 505
Db   1906 AGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTC 1965

Qy    506 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 525
Db   1966 CGCAACCCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAA 2025

Qy    526 SerAlaValGlnAlaAlaLeuGluLys 534
```

```
              :::   |||||||||||||||:::|||
Db     2026 GCCGTGGTTCAGGCCGCTCTGGATAAG 2052

RESULT 14
US-09-769-097-1
; Sequence 1, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
;  APPLICANT: Kimberly Anne Brun
;  APPLICANT:  Richard James Chenery
;  APPLICANT:  Harma Ellens
;  APPLICANT:  John Anthony Feild
;  APPLICANT:  Lin Yue
;  TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
;  TITLE OF INVENTION:  ENCODING RAT MDR1A AND SCREENING METHODS THEREOF
;  FILE REFERENCE: GP-50009-C2
;  CURRENT APPLICATION NUMBER: US/09/769,097
;  CURRENT FILING DATE:  2001-01-25
;  PRIOR APPLICATION NUMBER: 09/208,809
;  PRIOR FILING DATE: 1998-12-09
;  PRIOR APPLICATION NUMBER: 09/156,800
;  PRIOR FILING DATE: 1999-09-17
;  PRIOR APPLICATION NUMBER: US99/20770
;  PRIOR FILING DATE: 1999-09-10
;  NUMBER OF SEQ ID NOS: 4
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 4369
;   TYPE: DNA
;   ORGANISM: RATTUS RATTUS
US-09-769-097-1

Alignment Scores:
Pred. No.:               2.01e-157    Length:         4369
Score:                   1541.00      Matches:        300
Percent Similarity:      73.04%       Conservative:   101
Best Local Similarity:   54.64%       Mismatches:     102
Query Match:             56.26%       Indels:         46
DB:                      10           Gaps:           5

US-09-873-409-7 (1-541) x US-09-769-097-1 (1-4369)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
              |||:::||||||   ||||||:::::::::::|||      |||||||||||     |||
Db      499 ATGCTGCTGGGAACTCTGGCGGCCATTATCCATGGAATTGCGCTCCCACTTATGATGCTG 558

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
              |||   |||:::|||:::|||:::           :::      :::         :::
Db      559 GTCTTTGGAGACATGACAGATAGCTTTGCAAAT------GTAGGAAACAACCGTAGTATG 612

Qy       41 SerPhePheArgLeuThr------------------------------------------ 46
              ||||||:::      |||
Db      613 AGTTTCTACAATGCTACAGACATATATGCCAAGCTGGAGGACGAAATGGCCACGTACGCC 672

Qy       47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
                    ||||||   |||    |||            ||||||      |||||||||:::||||||
Db      673 TACTATTACACGGGCAATGGTGCCGGTGTGCTCATCGTTGCCTACATCCAGGTTTCACTT 732

Qy       67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
              |||      :::   |||   ||||||            :::||||||::::::|||||||||::::::
Db      733 TGGTGCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAAGTTTTTCCATGCCATC 792

Qy       87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
              :::      |||:::||||||||||||||||||         |||:::||||||||||||||||||:::
Db      793 ATGAATCAGGAGATAGGCTGGTTTGACGTGCATGACGTTGGGGAGCTCAACACCCGGCTC 852

Qy      107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
              |||      |||:::   ||||||::::::|||||||||||||||||      :::      ||||||
Db      853 ACAGATGACGTCTCCAAAATTAATGAAGGAATTGGTGACAAAATTGGAATGTTCTTTCAG 912

Qy      126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
                    |||:::||||||      |||         :::|||      :::||||||||||||||||
Db      913 GCAATGGCAACATTTTTTGGTGGTTTTATAATAGGATTTACTCGCGGCTGGAAGCTAACT 972

Qy      146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
              ||||||      |||:::      ||||||::::::      ||||||            :::::::::
Db      973 CTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTGTCAGCTGGTATTTGGGCAAAGATA 1032

Qy      166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
              :::   |||   |||   |||||||||   ||||||:::||||||||||||||||||
Db     1033 TTGTCTTCATTTACTGATAAGGAACTCCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAA 1092

Qy      186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
              ||||||||::::::||||||||||||||||||         |||:::|||||||||:::
Db     1093 GAAGTCTTAGCAGCCATCAGAACTGTGATTGCCTTTGGAGGACAAAAGAAGGAACTTGAA 1152

Qy      206 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeu 225

Db     1152 ------------------------------------------------------------ 1152

Qy      226 SerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 245
                                   ||||||      ||||||::::::||||||      |||
Db     1153 --------------------AGGTACAATAACAATTTGGAAGAAGCTAAAAGGCTTGGG 1191

Qy      246 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 265
              ||||||:::   |||   :::   :::|||:::||||||   :::      :::
Db     1192 ATAAAGAAAGCTATCACGGCCAACATTTCCATGGGTGCAGCTTTTCTGCTTATCTATGCA 1251

Qy      266 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 285
              :::|||   ||||||||||||||||||||||||||||||:::::::::   |||      |||
Db     1252 TCATATGCTCTGGCATTCTGGTATGGGACTTCCTTGGTCATCTCAAAAGAA------TAC 1305

Qy      286 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 305
              ||||||||   ||||||   ||||||||||||||||:::            ::::::      :::|||
Db     1306 ACTATTGGACAAGTGCTCACTGTCTTTTTTTCTGTATTAATTGGAGCATTCAGTGTTGGG 1365

Qy      306 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 325
                 |||   |||:::   |||   ||||||   ||||||||||||||||:::   :::|||
Db     1366 CAGGCATCTCCAAATATTGAAGCCTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTC 1425

Qy      326 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 345
                 :::||||||   ||||||||||||||||:::||||||   :::|||:::||||||:::
Db     1426 AGTATAATTGATAATAAGCCCAGTATAGACAGCTTCTCAAAGAGTGGGCACAAACCCGAC 1485

Qy      346 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 365
              :::|||:::|||   :::||||||||||||:::   |||:::||||||||||||
Db     1486 AACATACAAGGAAATTTGGAATTCAAAAATATTCACTTCAGTTACCCGTCTCGAAAAGAC 1545

Qy      366 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 385
              ::::::||||||||||||||||||||||||::::::||||||:::|||||||||||||||
Db     1546 GTTCAGATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTGGTT 1605

Qy      386 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 405
              |||      :::|||      ||||||||||||   |||||||||||||||||||||||||||
Db     1606 GGCAACAGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCC 1665

Qy      406 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 425
                 :::|||   :::   :::|||         ||||||||||||   :::|||||||||:::
Db     1666 ATAGAGGGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTG 1725

Qy      426 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 445
              |||:::   ||||||||||||||||||||||||||||||      |||||||||:::
Db     1726 CGGGAAATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCCACCACAATTGCCGAA 1785

Qy      446 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 465
              ||||||:::|||||||||::::::||||||   :::|||:::|||:::|||   :::|||
Db     1786 AACATTCGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAA 1845

Qy      466 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 485
              ||||||||||||||||||||||||:::   |||:::||||||:::|||||||||||||||
Db     1846 GCCAATGCCTATGATTTCATCATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAG 1905

Qy      486 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 505
              :::|||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db     1906 AGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTC 1965
```

```
Qy     506 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 525
Db    1966 CGCAACCCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAA 2025

Qy     526 SerAlaValGlnAlaAlaLeuGluLys 534
Db    2026 GCCGTGGTTCAGGCCGCTCTGGATAAG 2052
```

RESULT 15
US-10-072-621-2
; Sequence 2, Application US/10072621
; Patent No. US20020169137A1
; GENERAL INFORMATION:
; APPLICANT: Reiner, Peter B.
; APPLICANT: Connop, Bruce P.
; APPLICANT: Pollard, Michelle
; TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION
; TITLE OF INVENTION: BY MODIFICATION OF ABC TRANSPORTER EXPRESSION OR ACTIVITY
; FILE REFERENCE: 100103.402
; CURRENT APPLICATION NUMBER: US/10/072,621
; CURRENT FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 4643
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-10-072-621-2

```
Alignment Scores:
Pred. No.:              3.65e-157     Length:         4643
Score:                  1539.00       Matches:        305
Percent Similarity:     73.10%        Conservative:   100
Best Local Similarity:  55.05%        Mismatches:     99
Query Match:            56.19%        Indels:         50
DB:                     9             Gaps:           7
```

US-09-873-409-7 (1-541) x US-10-072-621-2 (1-4643)

```
Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db     575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy      21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
Db     635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db     695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db     755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db     815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db     875 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db     935 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAAATT 994

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db     995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy     201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db    1235 AAGAAAGAACTTGAA--------------------------------------------- 1249

Qy     221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db    1250 -----------------------------------AGGTACAACAAAAATTTAGAAGAA 1273

Qy     241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db    1274 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 1333

Qy     261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db    1334 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 1393

Qy     281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db    1394 GGGGAA------TATTCTATTGGACAAGTACTCACTGTA---TTTTCTGTATTAATTGGG 1444

Qy     301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db    1445 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1504

Qy     321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db    1505 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1564

Qy     341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db    1565 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1624

Qy     361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db    1625 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGTCTGAACCTGAAGGTGCAGAGTGGGCAG 1684

Qy     381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db    1685 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1744

Qy     401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db    1745 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1804

Qy     421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db    1805 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1864

Qy     441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db    1865 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1924

Qy     461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db    1925 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1984

Qy     481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db    1985 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 2044

Qy     501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db    2045 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 2104
```

```
Qy     521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
           |||:::||||||::::::   ||||||   ||||||:::|||
Db    2105 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 2146

Search completed: March 31, 2003, 16:14:26
Job time : 261.92 secs
```

```
                     GenCore version 5.1.4_p5_4578
              Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 30, 2003, 01:24:25 ; Search time 7313.73 Seconds
                                           (without alignments)
                                           2152.747 Million cell updates/sec

Title:          US-09-873-409-7
Perfect score:  2739
Sequence:       1 MILGILASLVNGACLPLMPL..........SESKSAVQAALEKDTPRYSF 541

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters:      4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115418_19228/app_query.fasta_1.7544
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_54887_@runat_27032003_115418_19228 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      GenEmbl:*
                1:  gb_ba:*
                2:  gb_htg:*
                3:  gb_in:*
                4:  gb_om:*
                5:  gb_ov:*
                6:  gb_pat:*
                7:  gb_ph:*
                8:  gb_pl:*
                9:  gb_pr:*
                10:  gb_ro:*
                11:  gb_sts:*
                12:  gb_sy:*
                13:  gb_un:*
                14:  gb_vi:*
                15:  em_ba:*
                16:  em_fun:*
                17:  em_hum:*
                18:  em_in:*
                19:  em_mu:*
                20:  em_om:*
                21:  em_or:*
                22:  em_ov:*
                23:  em_pat:*
                24:  em_ph:*
                25:  em_pl:*
                26:  em_ro:*
                27:  em_sts:*
                28:  em_un:*
                29:  em_vi:*
                30:  em_htg_hum:*
                31:  em_htg_inv:*
                32:  em_htg_other:*
                33:  em_htg_mus:*
                34:  em_htg_pln:*
                35:  em_htg_rod:*
                36:  em_htg_mam:*
                37:  em_htg_vrt:*
                38:  em_sy:*
                39:  em_htgo_hum:*
                40:  em_htgo_mus:*
                41:  em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2737	99.9	2021	6	AX339033	AX339033 Sequence
2	2697	98.5	3702	6	AX339031	AX339031 Sequence
3	2568.5	93.8	1940	6	AX339034	AX339034 Sequence
4	2528.5	92.3	3621	6	AX339032	AX339032 Sequence
5	2437	89.0	3699	6	AX478104	AX478104 Sequence
6	1868	68.2	3177	6	AX339030	AX339030 Sequence
7	1566.5	57.2	4189	6	AX322791	AX322791 Sequence
8	1566.5	57.2	4298	10	MUSMDR	M14757 Mouse multi
9	1561.5	57.0	4279	6	AX105082	AX105082 Sequence
10	1557.5	56.9	4279	6	AX105078	AX105078 Sequence
11	1557.5	56.9	4279	6	AX105080	AX105080 Sequence
12	1557	56.8	4195	6	AX108656	AX108656 Sequence
13	1556.5	56.8	4045	12	AF269224	AF269224 Synthetic
14	1555.5	56.8	3860	6	AX322787	AX322787 Sequence
15	1555.5	56.8	4378	6	E02326	E02326 Multidrug r
16	1555.5	56.8	4646	6	AX391099	AX391099 Sequence
17	1555	56.8	4280	10	CRUPGPII	M60041 C.griseus P
18	1552.5	56.7	3888	10	AY082609	AY082609 Rattus no
19	1552.5	56.7	4233	6	AR123273	AR123273 Sequence
20	1552.5	56.7	4279	6	AX105057	AX105057 Sequence
21	1552.5	56.7	4317	4	AF045016	AF045016 Canis fam
22	1552.5	56.7	4317	6	AX105059	AX105059 Sequence
23	1552.5	56.7	4669	6	AR091275	AR091275 Sequence
24	1552.5	56.7	4669	6	AR203322	AR203322 Sequence
25	1548.5	56.5	4186	6	AX108654	AX108654 Sequence
26	1548	56.5	4254	6	AX401748	AX401748 Sequence
27	1548	56.5	4254	10	RATMDRM	M81855 Rat mdr mRN
28	1548	56.5	4927	10	AF257746	AF257746 Rattus no
29	1547	56.5	4323	10	AF286167	AF286167 Rattus no
30	1546.5	56.5	3860	6	AX322789	AX322789 Sequence
31	1546.5	56.5	4646	6	AX336420	AX336420 Sequence
32	1546.5	56.5	4646	6	AX336708	AX336708 Sequence
33	1546.5	56.5	4646	6	I49610	I49610 Sequence 2
34	1546.5	56.5	4646	9	HUMMDR1	M14758 Homo sapien
35	1546.5	56.5	4669	6	I08557	I08557 Sequence 3
36	1546.5	56.5	6505	6	AR028671	AR028671 Sequence
37	1546.5	56.5	8630	6	AX012320	AX012320 Sequence
38	1546.5	56.5	8630	6	AX012321	AX012321 Sequence
39	1546.5	56.5	9318	6	AR028672	AR028672 Sequence
40	1545	56.4	4264	6	AR051647	AR051647 Sequence
41	1545	56.4	4264	6	AR051650	AR051650 Sequence
42	1542	56.3	4356	10	MUSMDRAA	M30697 Mouse multi
43	1541.5	56.3	4296	10	CRUPGPI	M60040 C.griseus P
44	1540.5	56.2	4304	10	CRUPGP1165	M59253 Chinese ham
45	1540	56.2	3489	4	AB029153	AB029153 Felis cat

ALIGNMENTS

RESULT 1

```
AX339033
LOCUS       AX339033                2021 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 15 from Patent WO0194400.
ACCESSION   AX339033
VERSION     AX339033.1  GI:18129125
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 15 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..2021
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      547 a    406 c    462 g    605 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              2.61e-214     Length:         2021
Score:                  2737.00       Matches:        541
Percent Similarity:     100.00%       Conservative:   0
Best Local Similarity:  100.00%       Mismatches:     0
Query Match:            99.93%        Indels:         0
DB:                     6             Gaps:           0

US-09-873-409-7 (1-541) x AX339033 (1-2021)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      34 ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 93

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153

Qy      41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     154 TCTTTCTTCAGGTTGACCCTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGT 213

Qy      61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     214 TACATACAGATTTCCTTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAA 273

Qy      81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     274 CAGTTTTTTCATTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333

Qy     101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     334 GAACTTAACACTCGCATGACAGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATT 393

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAG 453

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     454 GGCTGGAAACTCACCCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA 513

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAG 633

Qy     201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     634 GAGAAAGAACTTCAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTTGGTTTTATTTT 693

Qy     221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     694 CCCCAGTGGCTACTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGAATCTCAAAGAT 753

Qy     241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     754 GCAAAGGATTTTGGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTAC 813

Qy     261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     814 TTCTTTATGAATGGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAAT 873

Qy     281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     874 GGAGAACCTGGATATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGC 933

Qy     301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     934 AGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCT 993

Qy     321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     994 GCCTTTCATATTTTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCT 1053

Qy     341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1054 GGATATAAACCTGAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTAT 1113

Qy     361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1114 CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAG 1173

Qy     381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1174 ACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG 1233

Qy     401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1234 AGGTTATATGATCCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1293

Qy     421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1294 AATGTGCGGCATTATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGG 1353

Qy     441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1354 ACCACCATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1413

Qy     461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1414 AGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAAT 1473

Qy     481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1474 ACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATT 1533

Qy     501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1534 GCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTG 1593

Qy     521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAspThrProArgTyrSer 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1594 GATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTGGAGAAGGATACCCCCAGGTATTCA 1653

Qy     541 Phe 541
           |||
Db    1654 TTT 1656

RESULT 2
```

```
AX339031
LOCUS       AX339031                3702 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 13 from Patent WO0194400.
ACCESSION   AX339031
VERSION     AX339031.1  GI:18129123
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 13 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..3702
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1098 a    708 c    856 g   1039 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              1.04e-210      Length:         3702
Score:                  2697.00        Matches:        534
Percent Similarity:     100.00%        Conservative:   0
Best Local Similarity:  100.00%        Mismatches:     0
Query Match:            98.47%         Indels:         0
DB:                     6              Gaps:           0

US-09-873-409-7 (1-541) x AX339031 (1-3702)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      34 ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 93

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153

Qy      41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     154 TCTTTCTTCAGGTTGACCCTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGT 213

Qy      61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     214 TACATACAGATTTCCTTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAA 273

Qy      81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     274 CAGTTTTTTCATTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333

Qy     101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     334 GAACTTAACACTCGCATGACAGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATT 393

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAG 453

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     454 GGCTGGAAACTCACCCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA 513

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAG 633

Qy     201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     634 GAGAAAGAACTTCAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTTGGTTTTATTTT 693

Qy     221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     694 CCCCAGTGGCTACTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGAATCTCAAAGAT 753

Qy     241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     754 GCAAAGGATTTTGGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTAC 813

Qy     261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     814 TTCTTTATGAATGGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAAT 873

Qy     281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     874 GGAGAACCTGGATATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGC 933

Qy     301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     934 AGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCT 993

Qy     321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     994 GCCTTTCATATTTTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCT 1053

Qy     341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1054 GGATATAAACCTGAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTAT 1113

Qy     361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1114 CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAG 1173

Qy     381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1174 ACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG 1233

Qy     401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1234 AGGTTATATGATCCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1293

Qy     421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1294 AATGTGCGGCATTATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGG 1353

Qy     441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1354 ACCACCATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1413

Qy     461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1414 AGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAAT 1473

Qy     481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1474 ACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATT 1533

Qy     501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1534 GCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTG 1593

Qy     521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
             ||||||||||||||||||||||||||||||||||||||||||
Db    1594 GATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTGGAGAAG 1635

RESULT 3
AX339034
LOCUS       AX339034                1940 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 16 from Patent WO0194400.
ACCESSION   AX339034
```

```
VERSION     AX339034.1  GI:18129126
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 16 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..1940
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      530 a    394 c    448 g    568 t
ORIGIN

Alignment Scores:
Pred. No.:              1.45e-200     Length:         1940
Score:                  2568.50       Matches:        514
Percent Similarity:     95.01%        Conservative:   0
Best Local Similarity:  95.01%        Mismatches:     0
Query Match:            93.78%        Indels:         27
DB:                     6             Gaps:           1

US-09-873-409-7 (1-541) x AX339034 (1-1940)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     34 ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 93

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153

Qy     41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    154 TCTTTCTTCAGGTTGACCCTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGT 213

Qy     61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    214 TACATACAGATTTCCTTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAA 273

Qy     81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    274 CAGTTTTTTCATTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333

Qy    101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    334 GAACTTAACACTCGCATGACAGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATT 393

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAG 453

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    454 GGCTGGAAACTCACCCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA 513

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAG 633

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
          |||||||||||||||
Db    634 GAGAAAGAACTTCAA--------------------------------------------- 648

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                              ||||||||||||||||||||||||
Db    649 ------------------------------------AGGTATACACAGAATCTCAAAGAT 672

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    673 GCAAAGGATTTTGGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTAC 732

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    733 TTCTTTATGAATGGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAAT 792

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    793 GGAGAACCTGGATATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGC 852

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    853 AGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCT 912

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    913 GCCTTTCATATTTTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCT 972

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    973 GGATATAAACCTGAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTAT 1032

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1033 CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAG 1092

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1093 ACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG 1152

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1153 AGGTTATATGATCCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1212

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1213 AATGTGCGGCATTATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGG 1272

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1273 ACCACCATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1332

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1333 AGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAAT 1392

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1393 ACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATT 1452

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1453 GCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTG 1512

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAspThrProArgTyrSer 540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1513 GATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTGGAGAAGGATACCCCCAGGTATTCA 1572

Qy    541 Phe 541
          |||
Db   1573 TTT 1575

RESULT 4
AX339032
LOCUS       AX339032                3621 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 14 from Patent WO0194400.
ACCESSION   AX339032
```

```
VERSION     AX339032.1  GI:18129124
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 14 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..3621
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1081 a    696 c    842 g   1002 t
ORIGIN

Alignment Scores:
Pred. No.:               5.89e-197      Length:         3621
Score:                   2528.50        Matches:        507
Percent Similarity:      94.94%         Conservative:   0
Best Local Similarity:   94.94%         Mismatches:     0
Query Match:             92.31%         Indels:         27
DB:                      6              Gaps:           1

US-09-873-409-7 (1-541) x AX339032 (1-3621)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      34 ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 93

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153

Qy      41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     154 TCTTTCTTCAGGTTGACCCTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGT 213

Qy      61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     214 TACATACAGATTTCCTTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAA 273

Qy      81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     274 CAGTTTTTTCATTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333

Qy     101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     334 GAACTTAACACTCGCATGACAGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATT 393

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAG 453

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     454 GGCTGGAAACTCACCCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA 513

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAG 633

Qy     201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
           ||||||||||||||
Db     634 GAGAAAGAACTTCAA--------------------------------------------- 648

Qy     221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                                  |||||||||||||||||||||
Db     649 ------------------------------------AGGTATACACAGAATCTCAAAGAT 672

Qy     241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     673 GCAAAGGATTTTGGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTAC 732

Qy     261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     733 TTCTTTATGAATGGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAAT 792

Qy     281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     793 GGAGAACCTGGATATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGC 852

Qy     301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     853 AGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCT 912

Qy     321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     913 GCCTTTCATATTTTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCT 972

Qy     341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     973 GGATATAAACCTGAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTAT 1032

Qy     361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1033 CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAG 1092

Qy     381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1093 ACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG 1152

Qy     401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1153 AGGTTATATGATCCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1212

Qy     421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1213 AATGTGCGGCATTATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGG 1272

Qy     441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1273 ACCACCATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1332

Qy     461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1333 AGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAAT 1392

Qy     481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1393 ACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATT 1452

Qy     501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1453 GCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTG 1512

Qy     521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
           ||||||||||||||||||||||||||||||||||||||||||
Db    1513 GATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTGGAGAAG 1554

RESULT 5
AX478104
LOCUS       AX478104                3699 bp    DNA     linear   PAT 12-AUG-2002
DEFINITION  Sequence 34 from Patent WO0240541.
ACCESSION   AX478104
VERSION     AX478104.1  GI:22217064
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Tang,Y.T., Yue,H., Nguyen,D.B., Hafalia,A.J., Elliott,V.S., Lu,Y.,
            Walia,N.K., Yao,M.G., Baughn,M.R., Gandhi,A.R., Ding,L.,
            Sanjanwala,M., Ramkumar,J., Arvizu,C., Gietzen,K.J., Lal,P.G.,
            Azimzai,Y., Khan,F.A., Thangavelu,K., Thornton,M., Lu,D.A.,
            Tribouley,C.M., Warren,B.A., Ison,C.H., Das,D., Raumann,B.E.,
            Policky,J.L. and Kearney,L.
  TITLE     Transporters and ion channels
  JOURNAL   Patent: WO 0240541-A 34 23-MAY-2002;
            Incyte Genomics, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..3699
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 7472030CB1"
BASE COUNT     1116 a    707 c    860 g   1016 t
ORIGIN

Alignment Scores:
Pred. No.:              1.81e-189     Length:        3699
Score:                  2437.00       Matches:       495
Percent Similarity:     90.71%        Conservative:  3
Best Local Similarity:  90.16%        Mismatches:    9
Query Match:            88.97%        Indels:        42
DB:                     6             Gaps:          3

US-09-873-409-7 (1-541) x AX478104 (1-3699)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          ||||||||||||||||||    ||||||   ||||||||||||||||||||||||||||||
Db    145 ATGATCCTGGGTATACTGACATCACTGTTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 204

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
             :::||||||||||||||||||||||||||||||||||||||||    |||||||||
Db    205 TGTATAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCACACTAACACAACA 264

Qy     41 SerPhePheArg------------------------------------------LeuThr 46
          ::::::                                                ||||||
Db    265 AATTATCAGAACTGTACTCAGTCTCAAGAGAAGCTGAATGAAGATATGACTCTGTTGACC 324

Qy     47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    325 CTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGTTACATACAGATTTCCTTG 384

Qy     67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    385 TGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAACAGTTTTTTCATTCAGTT 444

Qy     87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    445 TTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGTGAACTTAACACTCGCATG 504

Qy    107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
          |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    505 ACAGATGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATTGCTCTGTTGTTTCAA 564

Qy    126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    565 AACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAGGGCTGGAAACTCACC 624

Qy    146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    625 CTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCAGCATGTTCTAGGATG 684

Qy    166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    685 GTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCAGAA 744

Qy    186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    745 GAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAGGAGAAAGAACTTCAA 804

Qy    206 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeu 225

Db    804 ------------------------------------------------------------ 804

Qy    226 SerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 245
                             |||||||||||||||||||||||||||||||||||||||
Db    805 ---------------------AGGTATACACAGAATCTCAAAGATGCAAAGGATTTTGGC 843

Qy    246 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 265
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    844 ATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAATGGA 903

Qy    266 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 285
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    904 ACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGATAT 963

Qy    286 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    964 ACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGTAGTTATTGCATTGGA 1023

Qy    306 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 325
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1024 GCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTC 1083

Qy    326 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 345
          |||||||||||||||||||||||||||    |||||||||||||||||||||||||||||
Db   1084 CAGGTTATTGATAAGAAACCCAGTATAGGTAACTTTTCCACAGCTGGATATAAACCTGAA 1143

Qy    346 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 365
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1144 TCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGGCCATCT 1203

Qy    366 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 385
          ||||||||||||||||||||||||||   |||||||||||||||||||||||||||||||
Db   1204 ATCAAGATTCTGAAAGGTCTGAATCTCGGAATTAAGTCTGGAGAGACAGTCGCCTTGGTC 1263

Qy    386 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 405
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1264 GGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCG 1323

Qy    406 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 425
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1324 GATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTAT 1383

Qy    426 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 445
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1384 CGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAAC 1443

Qy    446 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 465
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1444 AATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAA 1503

Qy    466 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 485
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1504 GCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAA 1563

Qy    486 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 505
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1564 AAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTT 1623

Qy    506 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 525
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1624 CGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAG 1683

Qy    526 SerAlaValGlnAlaAlaLeuGluLys 534
          |||||||||||||||||||||||||||
Db   1684 TCAGCTGTTCAAGCTGCACTGGAGAAG 1710

RESULT 6
AX339030
LOCUS       AX339030        3177 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 12 from Patent WO0194400.
```

```
ACCESSION   AX339030
VERSION     AX339030.1  GI:18129122
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 12 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1. .3177
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      970 a    601 c    735 g    870 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              4.58e-143     Length:        3177
Score:                  1868.00       Matches:       370
Percent Similarity:     100.00%       Conservative:  0
Best Local Similarity:  100.00%      .Mismatches:    0
Query Match:            68.20%        Indels:        0
DB:                     6             Gaps:          0

US-09-873-409-7 (1-541) x AX339030 (1-3177)

Qy     165 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 184
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCA 60

Qy     185 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 204
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAGGAGAAAGAACTT 120

Qy     205 GlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeu 224
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTTGGTTTTATTTCCCCAGTGGCTA 180

Qy     225 LeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 244
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 CTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGAATCTCAAAGATGCAAAGGATTTT 240

Qy     245 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 264
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAAT 300

Qy     265 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 284
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGA 360

Qy     285 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 304
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATT 420

Qy     305 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 324
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATT 480

Qy     325 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 344
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCT 540

Qy     345 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 364
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCA 600

Qy     365 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 384
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 TCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTG 660

Qy     385 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 404
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 720

Qy     405 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 424
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 CCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCAT 780

Qy     425 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 444
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 TATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGT 840

Qy     445 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 464
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 AACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGG 900

Qy     465 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 484
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGG 960

Qy     485 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 504
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTA 1020

Qy     505 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 524
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGC 1080

Qy     525 LysSerAlaValGlnAlaAlaLeuGluLys 534
           ||||||||||||||||||||||||||||||
Db    1081 AAGTCAGCTGTTCAAGCTGCACTGGAGAAG 1110

RESULT 7
AX322791
LOCUS       AX322791                4189 bp    DNA     linear   PAT 07-JAN-2002
DEFINITION  Sequence 5 from Patent WO0192877.
ACCESSION   AX322791
VERSION     AX322791.1  GI:18093768
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Sorrentino,B. and Schuetz,J.
  TITLE     Method of identifying and/or isolating stem cells
  JOURNAL   Patent: WO 0192877-A 5 06-DEC-2001;
            ST. JUDE CHILDREN'S RESEARCH HOSPITAL (US)
FEATURES             Location/Qualifiers
     source          1. .4189
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
BASE COUNT     1204 a    875 c   1028 g   1082 t
ORIGIN

Alignment Scores:
Pred. No.:              2.78e-118     Length:        4189
Score:                  1566.50       Matches:       307
Percent Similarity:     72.63%        Conservative:  99
Best Local Similarity:  54.92%        Mismatches:    94
Query Match:            57.19%        Indels:        59
DB:                     6             Gaps:          6

US-09-873-409-7 (1-541) x AX322791 (1-4189)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||||||||||   |||||::::::::::::|||         |||||||||:::    |||
Db     148 ATGATTCTGGGAACTCTCGCTGCTATTATCCATGGAACATTACTTCCCCTCTTGATGCTG 207

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
           |||   |||    |||:::|||:::
Db     208 GTGTTTGGAAACATGACAGATAGTTTTACAAAAGCAGAAGCCAGTATTCTGCCAAGCATT 267
```

```
Qy     31 -----------------SerGlyCysLeuValGlnThrAsnThr-------------- 39
                           :::      |||:::    :::|||:::
Db    268 ACTAATCAAAGTGGACCCAACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAAGAG 327

Qy     40 ---------TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56
                   |||::::::                   ||||||   ||||||||
Db    328 ATGGCCATATACGCCTAC---------------TATTACACCGGGATTGGTGCTGGTGTG 372

Qy     57 LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys 76
          ||||||       |||||||||:::|||||||||    :::    |||    ||||||
Db    373 CTCATAGTTGCCTACATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGAAGACAGATACAC 432

Qy     77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96
          :::||||||::::::|||||||||:::::::::     |||:::||||||||||||||
Db    433 AAGATTAGGCAGAAGTTTTTCCATGCTATAATGAATCAGGAGATAGGCTGGTTTGATGTG 492

Qy     97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
             |||:::||||||||||||||||||:::|||      |||:::    ||||||:::||||||
Db    493 CATGATGTTGGGGAGCTCAACACCCGGCTCACAGATGATGTCTCCAAAATTAATGACGGA 552

Qy    116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
          |||||||||||||||     :::     ||||||:::::::::||||||       |||
Db    553 ATTGGTGACAAAATTGGGATGTTTTTTCAGTCCATAACCACATTTTTAGCCGGTTTTATC 612

Qy    136 ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
          :::|||     :::    |||||||||||||||||||||      |||:::    ||||||||||||
Db    613 ATAGGATTTATAAGTGGTTGGAAGCTAACCCTTGTCATTTTGGCTGTCAGCCCTCTTATT 672

Qy    156 MetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
                |||:::|||       :::::::::::     |||      |||:::|||||||||
Db    673 GGATTGTCATCTGCTTTGTGGGCAAAGGTATTGACTTCATTTACTAATAAGGAACTCCAG 732

Qy    176 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIle 195
          ||||||:::||||||||||||||||||||||||||||||||:::::::||||||||||||
Db    733 GCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT 792

Qy    196 AlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyr 215
          ||||||       |||:::|||||||||:::
Db    793 GCCTTTGGAGGACAACAGAAGGAACTTGAA---------------------------- 822

Qy    216 AlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThr 235
                                                            ||||||
Db    823 --------------------------------------------------AGGTACAAT 831

Qy    236 GlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSer 255
          :::||||||::::::||||||:::      ||||||||||:::    |||     :::    :::|||
Db    832 AAAAATTTAGAAGAAGCTAAAAATGTTGGCATAAAGAAAGCTATCACAGCCAGCATTTCG 891

Qy    256 LeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThr 275
          :::|||       |||         :::         :::|||    |||||||||||||||||||
Db    892 ATAGGCATTGCCTACCTGTTGGTCTATGCATCATATGCACTGGCATTCTGGTATGGGACA 951

Qy    276 SerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePhe 295
          ||||||:::|||:::      |||           |||:::||||||      ||||||    |||||||||
Db    952 TCCTTGGTCCTCTCAAATGAA------TATTCTATTGGAGAAGTGCTTACTGTCTTCTTC 1005

Qy    296 SerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAla 315
          |||::::::         ::::::    ||||||               |||:::    |||    ||||||
Db   1006 TCTATTTTGTTGGGGACTTTTAGTATTGGACACTTGGCCCCAAACATAGAAGCCTTTGCA 1065

Qy    316 IleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAsp 335
             ||||||||||||||||||   ||||||::::::||||||     :::||||||||||||
Db   1066 AACGCACGAGGGGCAGCCTTTGAAATCTTCAAGATAATTGATAACGAGCCAAGCATTGAC 1125

Qy    336 AsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsn 355
          :::||||||||      |||||||||||:::||||||      |||         :::||||||||||||
Db   1126 AGCTTCTCAACAAAGGGCTACAAACCAGACAGTATAATGGGAAACTTAGAGTTTAAAAAT 1185

Qy    356 ValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArg 375
          |||      |||||||||||||||||            ::::::||||||||||||||||||||:::
Db   1186 GTTCACTTCAACTACCCATCGAGAAGCGAAGTTCAGATCTTGAAGGGCCTCAATCTGAAG 1245

Qy    376 IleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVal 395
          :::|||||||||:::|||||||||||||||||||       :::|||      |||||||||||
Db   1246 GTGAAGAGCGGACAGACGGTGGCCTTGGTTGGCAACAGTGGCTGTGGAAAAAGCACAACT 1305

Qy    396 ValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsn 415
          |||||||||:::|||||||||||||||||||       :::|||    :::      :::|||
Db   1306 GTCCAGCTGATGCAGAGGCTCTACGACCCCCTGGAGGGCGTGGTCAGTATCGACGGACAA 1365

Qy    416 AspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGlu 435
          ||||||||    :::|||||||||:::    |||:::    ||||||||||||||||||||||
Db   1366 GACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCATTGGTGTGGTGAGTCAGGAA 1425

Qy    436 ProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThr 455
          ||||||||||||    ||||||||||:::    ||||||:::|||||||||:::||||||||
Db   1426 CCTGTGCTGTTTGCCACCACGATCGCCGAGAACATTCGCTATGGCCGAGAAGATGTCACC 1485

Qy    456 AspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhe 475
             :::|||:::|||:::|||    :::||||||||||||||||||||||||||||||:::
Db   1486 ATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTCATCATGAAACTG 1545

Qy    476 ProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLys 495
          |||::::::|||:::|||||||||||||||||||||:::|||||||||:::||||||||||||
Db   1546 CCCCACCAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAA 1605

Qy    496 GlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGlu 515
          |||||||||||||||||||||||||||||||||||||||||||||:::||||||||||
Db   1606 CAGAGAATCGCCATTGCCCGGGCCCTGGTCCGCAATCCCAAGATCCTTTTGTTGGACGAG 1665

Qy    516 AlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
          ||||||||||||||||:::||||||::::::    |||||||||||||:::|||
Db   1666 GCCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTGGTGCAGGCCGCACTGGATAAG 1722
```

```
RESULT 8
MUSMDR
LOCUS       MUSMDR                 4298 bp    mRNA    linear   ROD 04-OCT-1995
DEFINITION  Mouse multidrug resistant protein mRNA, complete cds.
ACCESSION   M14757
VERSION     M14757.1  GI:199100
KEYWORDS    membrane-associated glycoprotein; multidrug resistance protein.
SOURCE      Mus musculus (strain BALB/c, sub_species domesticus) (clone: lambda
            DR11) cDNA to mRNA.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4298)
  AUTHORS   Gros,P., Croop,J. and Housman,D.
  TITLE     Mammalian multidrug resistance gene: complete cDNA sequence
            indicates strong homology to bacterial transport proteins
  JOURNAL   Cell 47 (3), 371-380 (1986)
  MEDLINE   87028229
   PUBMED   3768958
COMMENT     The mdr gene family includes two or more related but distinct
            cellular genes.
            There are 3 possible start codons besides the one given in the
            Features table, at positions 161-163, 221-223 and 257-259.  If the
            latter two start codons are used as initiator the first hydrophobic
            domain could possibly serve as a signal sequence.
FEATURES             Location/Qualifiers
     source          1..4298
                     /organism="Mus musculus"
                     /strain="BALB/c"
                     /sub_species="domesticus"
                     /db_xref="taxon:10090"
                     /clone="lambda DR11"
                     /cell_type="pre-B cell"
     mRNA            <1..4298
                     /product="mdr mRNA"
     CDS             110..3940
                     /codon_start=1
                     /product="multidrug resistance protein"
                     /protein_id="AAA79005.1"
                     /db_xref="GI:387426"
                     /translation="MEFEENLKGRADKNFSKMGKKSKKEKKEKKPAVGVFGMFRYADW
```

```
                 LDKLCMILGTLAAIIHGTLLPLLMLVFGNMTDSFTKAEASILPSITNQSGPNSTLIIS
                 NSSLEEEMAIYAYYYTGIGAGVLIVAYIQVSLWCLAAGRQIHKIRQKFFHAIMNQEIG
                 WFDVHDVGELNTRLTDDVSKINDGIGDKIGMFFQSITTFLAGFIIGFISGWKLTLVIL
                 AVSPLIGLSSALWAKVLTSFTNKELQAYAKAGAVAEEVLAAIRTVIAFGGQQKELERY
                 NKNLEEAKNVGIKKAITASISIGIAYLLVYASYALAFWYGTSLVLSNEYSIGEVLTVF
                 FSILLGTFSIGHLAPNIEAFANARGAAFEIFKIIDNEPSIDSFSTKGYKPDSIMGNLE
                 FKNVHFNYPSRSEVQILKGLNLKVKSGQTVALVGNSGCGKSTTVQLMQRLYDPLEGVV
                 SIDGQDIRTINVRYLREIIGVVSQEPVLFATTIAENIRYGREDVTMDEIEKAVKEANA
                 YDFIMKLPHQFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTESEA
                 VVQAALDKAREGRTTIVIAHRLSTVRNADVIAGFDGGVIVEQGNHDELMREKGIYFKL
                 VMTQTRGNEIEPGNNAYGSQSDTDASELTSEESKSPLIRRSIYRSVHRKQDQERRLSM
                 KEAVDEDVPLVSFWRILNLNLSEWPYLLVGVLCAVINGCIQPVFAIVFSRIVGVFSRD
                 DDHETKRQNCNLFSLFFLVMGLISFVTYFFQGFTFGKAGEILTKRVRYMVFKSMLRQD
                 ISWFDDHKNSTGSLTTRLASDASSVKGAMGARLAVVTQNVANLGTGVILSLVYGWQLT
                 LLLVVIIPLIVLGGIIEMKLLSGQALKDKKQLEISGKIATEAIENFRTIVSLTREQKF
                 ETMYAQSLQVPYRNAMKKAHVFGITFSFTQAMMYFSYAACFRFGAYLVAQQLMTFENV
                 MLVFSAVVFGAMAAGNTSSFAPDYAKAKVSASHIIRIIEKTPEIDSYSTEGLKPTLLE
                 GNVKFNGVQFNYPTRPNIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERFYDPM
                 AGSVFLDGKEIKQLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRAVSHEEIVRA
                 AKEANIHQFIDSLPDKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEATSAL
                 DTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVIENGKVKEHGTHQQLLAQK
                 GIYFSMVQAGAKRS"
BASE COUNT     1226 a    903 c    1059 g    1110 t
ORIGIN

Alignment Scores:
Pred. No.:              2.88e-118      Length:         4298
Score:                  1566.50        Matches:        307
Percent Similarity:     72.63%         Conservative:   99
Best Local Similarity:  54.92%         Mismatches:     94
Query Match:            57.19%         Indels:         59
DB:                     10             Gaps:           6

US-09-873-409-7 (1-541) x MUSMDR (1-4298)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    257 ATGATTCTGGGAACTCTCGCTGCTATTATCCATGGAACATTACTTCCCCTCTTGATGCTG 316

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
Db    317 GTGTTTGGAAACATGACAGATAGTTTTACAAAAGCAGAAGCCAGTATTCTGCCAAGCATT 376

Qy     31 ------------------SerGlyCysLeuValGlnThrAsnThr--------------- 39
Db    377 ACTAATCAAAGTGGACCCAACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAAGAG 436

Qy     40 ---------TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56
Db    437 ATGGCCATATACGCCTAC---------------TATTACACCGGGATTGGTGCTGGTGTG 481

Qy     57 LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys 76
Db    482 CTCATAGTTGCCTACATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGAAGACAGATACAC 541

Qy     77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96
Db    542 AAGATTAGGCAGAAGTTTTTCCATGCTATAATGAATCAGGAGATAGGCTGGTTTGATGTG 601

Qy     97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
Db    602 CATGATGTTGGGGAGCTCAACACCCGGCTCACAGATGATGTCTCCAAAATTAATGACGGA 661

Qy    116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
Db    662 ATTGGTGACAAAATTGGGATGTTTTTTCAGTCCATAACCACATTTTTAGCCGGTTTTATC 721

Qy    136 ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
Db    722 ATAGGATTTATAAGTGGTTGGAAGCTAACCCTTGTCATTTTGGCTGTCAGCCCTCTTATT 781

Qy    156 MetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
Db    782 GGATTGTCATCTGCTTTGTGGGCAAAGGTATTGACTTCATTTACTAATAAGGAACTCCAG 841

Qy    176 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIle 195
Db    842 GCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT 901

Qy    196 AlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyr 215
Db    902 GCCTTTGGAGGACAACAGAAGGAACTTGAA------------------------------ 931

Qy    216 AlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThr 235
Db    932 ---------------------------------------------------AGGTACAAT 940

Qy    236 GlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSer 255
Db    941 AAAAATTTAGAAGAAGCTAAAAATGTTGGCATAAAGAAAGCTATCACAGCCAGCATTTCG 1000

Qy    256 LeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThr 275
Db   1001 ATAGGCATTGCCTACCTGTTGGTCTATGCATCATATGCACTGGCATTCTGGTATGGACA 1060

Qy    276 SerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePhe 295
Db   1061 TCCTTGGTCCTCTCAAATGAA------TATTCTATTGGAGAAGTGCTTACTGTCTTCTTC 1114

Qy    296 SerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAla 315
Db   1115 TCTATTTTGTTGGGGACTTTTAGTATTGGACACTTGGCCCCAAACATAGAAGCCTTTGCA 1174

Qy    316 IleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAsp 335
Db   1175 AACGCACGAGGGGCAGCCTTTGAAATCTTCAAGATAATTGATAACGAGCCAAGCATTGAC 1234

Qy    336 AsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsn 355
Db   1235 AGCTTCTCAACAAAGGGCTACAAACCAGACAGTATAATGGGAAACTTAGAGTTTAAAAAT 1294

Qy    356 ValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArg 375
Db   1295 GTTCACTTCAACTACCCATCGAGAAGCGAAGTTCAGATCTTGAAGGGCCTCAATCTGAAG 1354

Qy    376 IleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVal 395
Db   1355 GTGAAGAGCGGACAGACGGTGGCCTTGGTTGGCAACAGTGGCTGTGGAAAAAGCACAACT 1414

Qy    396 ValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsn 415
Db   1415 GTCCAGCTGATGCAGAGGCTCTACGACCCCCTGGAGGGCGTGGTCAGTATCGACGGACAA 1474

Qy    416 AspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGlu 435
Db   1475 GACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCATTGGTGTGGTGAGTCAGGAA 1534

Qy    436 ProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThr 455
Db   1535 CCTGTGCTGTTTGCCACCACGATCGCCGAGAACATTCGCTATGGCCGAGAAGATGTCACC 1594

Qy    456 AspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhe 475
Db   1595 ATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTCATCATGAAACTG 1654

Qy    476 ProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLys 495
Db   1655 CCCCACCAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAA 1714

Qy    496 GlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGlu 515
Db   1715 CAGAGAATCGCCATTGCCCGGGCCCTGGTCCGCAATCCCAAGATCCTTTTGTTGGACGAG 1774

Qy    516 AlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db   1775 GCCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTGGTGCAGGCCGCACTGGATAAG 1831
```

```
RESULT 9
AX105082
LOCUS       AX105082                4279 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 26 from Patent WO0123540.
ACCESSION   AX105082
VERSION     AX105082.1  GI:13921232
KEYWORDS    .
SOURCE      dog.
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 4279)
  AUTHORS   Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and
            Patten,C.J.
  TITLE     P-glycoproteins and uses thereof
  JOURNAL   Patent: WO 0123540-A 26 05-APR-2001;
            GENTEST CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..4279
                     /organism="Canis familiaris"
                     /db_xref="taxon:9615"
     CDS             17..3862
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC37723.1"
                     /db_xref="GI:13921233"
                     /translation="MDPEGGRKGSAEKNFWKMGKKSKKKEKKEKKPTVSTFAMFRYSN
                     WLDRLYMLVGTMAAIIHGAALPLMMLVFGNMTDSFANAGISRNKTFPVIINESITNNT
                     QHFINHLEEEMTTYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQILKIRKQFFHAIMRQ
                     EIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSIATFFTGFIVGFTRGWKLTL
                     VILAISPVLGLSAAIWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFGGQKKEL
                     ERYNKNLEEAKGIGIKKAITANISIGAAFLLIYASYALAFWYGTSLVLSSEYTIGQVL
                     TVFFSVLIGAFSIGQASPSIEAFANARGAAYEIFKIIDNKPSIDSYSKSGHKPDNIKG
                     NLEFKNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQRLYDPTD
                     GMVCIDGQDIRTINVRHLREITGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKE
                     ANAYDFIMKLPNKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTE
                     SEAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELMKEKGIY
                     FKLVTMQTRGNEIELENATGESKSESDALEMSPKDSGSSLIKRRSTRRSIHAPQGQDR
                     KLGTKEDLNENVPPVSFWRILKLNSTEWPYFVVGIFCAIINGGLQPAFSIIFSRIIGI
                     FTRDEDPETKRQNSNMFSVLFLVLGIISFITFFLQGFTFGKAGEILTKRLRYMVFRSM
                     LRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAVITQNIANLGTGIIISLIYG
                     WQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRTVVSLTR
                     EQKFEYMYAQSLQVPYRNSLRKAHIFGVSFSITQAMMYFSYAGCFRFGAYLVANEFMN
                     FQDVLLVFSAIVFGAMAVGQVSSFAPDYAKAKVSAAHVIMIIEKSPLIDSYSPHGLKP
                     NTLEGNVTFNEVVFNYPTRPDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERF
                     YDPLAGSVLIDGKEIKHLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSHEE
                     IVQAAKEANIHHFIETLPEKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEA
                     TSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGKVKEHGTHQQL
                     LAQKGIYFSMVSVQAGAKR"
BASE COUNT     1296 a    833 c   1009 g   1141 t
ORIGIN

Alignment Scores:
Pred. No.:              7.32e-118      Length:         4279
Score:                  1561.50        Matches:        304
Percent Similarity:     72.12%         Conservative:   97
Best Local Similarity:  54.68%         Mismatches:     104
Query Match:            57.01%         Indels:         51
DB:                     6              Gaps:           5

US-09-873-409-7 (1-541) x AX105082 (1-4279)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          |||:::::|||   :::|||:::::::::::||||||    |||||||||||   |||
Db    167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
          |||   |||   |||:::|||:::      :::      :::
Db    227 GTTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy     35 --------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                              :::      ||||||      |||
Db    287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy     45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                         ||||||   |||||||||          |||:::
Db    347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy     60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
             |||||||||:::|||   |||   :::   |||   ||||||      :::||||||
Db    407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy     80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
          ||||||||||||||||:::::::::      |||:::|||||||||||||||      |||:::
Db    467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy    100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
          ||||||||||||||||||:::|||   |||:::   ||||||:::::|||||||||||||
Db    527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy    119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
          ||||||   :::   ||||||:::::::::||||||      |||      ||||||
Db    587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy    139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
             :::||||||||||||||||||||||   |||:::   ||||||::::::      |||
Db    647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy    159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
          ||||||      ::::::::::::   |||   |||   |||||||||   ||||||:::
Db    707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy    179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
          ||||||||||||||||||||||||||||||::::::||||||||||||||||||||||||
Db    767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy    199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
             |||:::||||||||||:::
Db    827 GGACAAAAGAAAGAACTTGAA--------------------------------------- 847

Qy    219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
                                                      ||||||   :::||||||
Db    848 ---------------------------------------AGGTACAACAAAAATTTA 865

Qy    239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
          ::::::||||||      |||||||||:::   |||   :::   :::|||:::||||||
Db    866 GAAGAAGCTAAAGGAATTGGGATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCC 925

Qy    259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
             :::      :::      :::|||   ||||||||||||||||||||||||||:::
Db    926 GCTTTCTTATTGATCTATGCATCATATGCTCTGGCTTTCTGGTATGGACCTCCTTGGTC 985

Qy    279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
          |||:::   |||      ||||||||||||   ||||||   |||||||||||||||:::
Db    986 CTCTCCAGTGAA------TATACTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTA 1039

Qy    299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
                ::::::   ||||||   |||   |||         |||   ||||||   ||||||
Db   1040 ATTGGGGCTTTTAGTATTGGACAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGA 1099

Qy    319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
          |||||||||:::      |||||:::::::||||||   ||||||||||||||||::::::|||
Db   1100 GGAGCAGCTTATGAAATCTTCAAGATAATTGACAATAAACCAAGCATTGACAGCTATTCG 1159

Qy    339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
            :::|||:::|||||||:::::::|||:::|||   :::|||||||||||||||   |||
Db   1160 AAGAGTGGACATAAACCAGATAATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTC 1219

Qy    359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
          :::|||||||||||      :::|||||||||||||||||||||||||:::::::::|||
Db   1220 AGTTACCCTTCTCGAAAAGAAGTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGT 1279

Qy    379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
          |||:::||||||||||||||||||   :::|||   ||||||||||||   |||||||||
Db   1280 GGGCAGACAGTGGCGCTGGTTGGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTG 1339
```

```
Qy    399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
Db   1340 ATGCAGAGGCTCTATGACCCCACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGG 1399

Qy    419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db   1400 ACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTG 1459

Qy    439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
Db   1460 TTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAG 1519

Qy    459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
Db   1520 ATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTTATCATGAAACTACCTAATAAA 1579

Qy    479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db   1580 TTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATC 1639

Qy    499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db   1640 GCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCA 1699

Qy    519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db   1700 GCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCCCTGGATAAG 1747

RESULT 10
AX105078
LOCUS       AX105078                 4279 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 22 from Patent WO0123540.
ACCESSION   AX105078
VERSION     AX105078.1  GI:13921228
KEYWORDS    .
SOURCE      dog.
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 4279)
  AUTHORS   Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and
            Patten,C.J.
  TITLE     P-glycoproteins and uses thereof
  JOURNAL   Patent: WO 0123540-A 22 05-APR-2001;
            GENTEST CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..4279
                     /organism="Canis familiaris"
                     /db_xref="taxon:9615"
     CDS             17..3862
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC37721.1"
                     /db_xref="GI:13921229"
                     /translation="MDPEGGRKGSAEKNFWKMGKKSKKNEKKEKKPTVSTFAMFRYSN
                     WLDRLYMLVGTMAAIIHGAALPLMMLVFGNMTDSFANAGISRNKTFPVIINESITNNT
                     QHFINHLEEEMTTYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQILKIRKQFFHAIMRQ
                     EIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSIATFFTGFIVGFTRGWKLTL
                     VILAISPVLGLSAAIWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFGGQKKEL
                     ERYNKNLEEAKGIGIKKAITANISIGAAFLLIYASYALAFWYGTSLVLSSEYSIGQVL
                     TVFFSVLIGAFSIGQASPSIEAFANARGAAYEIFKIIDNKPSIDSYSKSGHKPDNIKG
                     NLEFKNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQRLYDPTD
                     GMVCIDGQDIRTINVRHLREITGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKE
                     ANAYDFIMKLPNKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTE
                     SEAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELMKEKGIY
                     FKLVTMQTRGNEIELENATGESKSESDALEMSPKDSGSSLIKRRSTRRSIHAPQGQDR
                     KLGTKEDLNENVPPVSFWRILKLNSTEWPYFVVGIFCAIINGGLQPAFSIIFSRIIGI
                     FTRDEDPETKRQNSNMFSVLFLVLGIISFITFFLQGFTFGKAGEILTKRLRYMVFRSM
                     LRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAVITQNIANLGTGIIISLIYG
                     WQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRTVVSLTR
                     EQKFEYMYAQSLQVPYRNSLRKAHIFGVSFSITQAMMYFSYAGCFRFGAYLVANEFMN
                     FQDVLLVFSAIVFGAMAVGQVSSFAPDYAKAKVSAAHVIMIIEKSPLIDSYSPHGLKP
                     NTLEGNVTFNEVVFNYPTRPDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERF
                     YDPLAGSVLIDGKEIKHLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSHEE
                     IMQAAKEANIHHFIETLPEKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEA
                     TSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGKVKEHGTHQQL
                     LAQKGIYFSMVSVQAGAKR"
BASE COUNT     1295 a    833 c   1008 g   1143 t
ORIGIN

Alignment Scores:
Pred. No.:              1.55e-117      Length:         4279
Score:                  1557.50        Matches:        303
Percent Similarity:     72.12%         Conservative:   98
Best Local Similarity:  54.50%         Mismatches:     104
Query Match:            56.86%         Indels:         51
DB:                     6              Gaps:           5

US-09-873-409-7 (1-541) x AX105078 (1-4279)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
Db    227 GTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy     35 --------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
Db    287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy     45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db    347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy     60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db    407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy     80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db    467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy    100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db    527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy    119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db    587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy    139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db    647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy    159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db    707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy    179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db    767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy    199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
Db    827 GGACAAAAGAAAGAACTTGAA--------------------------------------- 847

Qy    219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
Db    848 ----------------------------------------AGGTACAACAAAAATTTA 865

Qy    239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
Db    866 GAAGAAGCTAAAGGAATTGGGATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCC 925
```

```
Qy      259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
              :::       :::      :::|||   |||||||||||||||||||||||||||:::
Db      926 GCTTTCTTATTGATCTATGCATCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTC 985

Qy      279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
            |||:::   |||      |||:::|||||   ||||||   ||||||||||||||:::
Db      986 CTCTCCAGTGAA------TATTCTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTA 1039

Qy      299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
                  ::::::   ||||||   |||   |||      |||   ||||||   ||||||
Db     1040 ATTGGGGCTTTTAGTATTGGACAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGA 1099

Qy      319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
            |||||||||:::   ||||||::::::||||||   ||||||||||||||||::::::|||
Db     1100 GGAGCAGCTTATGAAATCTTCAAGATAATTGACAATAAACCAAGCATTGACAGCTATTCG 1159

Qy      339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
              :::|||:::||||||::::::|||:::|||   :::||||||||||||||||   |||
Db     1160 AAGAGTGGACATAAACCAGATAATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTC 1219

Qy      359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
              :::|||||||||||      :::||||||||||||||||||||||||:::::::::|||
Db     1220 AGTTACCCTTCTCGAAAAGAAGTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGT 1279

Qy      379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
            |||:::||||||||||||||||||   :::|||   ||||||||||||   |||||||||
Db     1280 GGCAGACAGTGGCGCTGGTTGGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTG 1339

Qy      399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
            :::||||||||||||||||||   ||||||   :::   :::|||      |||||||||
Db     1340 ATGCAGAGGCTCTATGACCCCACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGG 1399

Qy      419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
              :::||||||||||||   |||:::      ||||||||||||||||||||||||||||
Db     1400 ACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTG 1459

Qy      439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
            |||   ||||||||:::   |||:::|||||||::::::||||||   :::|||
Db     1460 TTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAG 1519

Qy      459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
            :::|||:::|||   :::||||||||||||||||||||||||||:::   |||||||||
Db     1520 ATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTTATCATGAAACTACCTAATAAA 1579

Qy      479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
            |||:::||||||||||||:::|||||||||:::||||||||||||||||||||||||
Db     1580 TTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATC 1639

Qy      499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
            ||||||||||||||||||||||||||||||||||:::||||||||||||||||||||||
Db     1640 GCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCA 1699

Qy      519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
            |||||||||:::||||||::::::   ||||||   ||||||:::|||
Db     1700 GCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCCCTGGATAAG 1747
```

```
RESULT 11
AX105080
LOCUS       AX105080                4279 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 24 from Patent WO0123540.
ACCESSION   AX105080
VERSION     AX105080.1  GI:13921230
KEYWORDS    .
SOURCE      dog.
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 4279)
  AUTHORS   Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and
            Patten,C.J.
  TITLE     P-glycoproteins and uses thereof
  JOURNAL   Patent: WO 0123540-A 24 05-APR-2001;
            GENTEST CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..4279
                     /organism="Canis familiaris"
                     /db_xref="taxon:9615"
     CDS             17..3862
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC37722.1"
                     /db_xref="GI:13921231"
                     /translation="MDPEGGRKGSAEKNFWKMGKKSKKKEKKEKKPTVSTFAMFRYSN
                     WLDRLYMLVGTMAAIIHGAALPLMMLVFGNMTDSFANAGISRNKTFPVIINESITNNT
                     QHFINHLEEEMTTYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQILKIRKQFFHAIMRQ
                     EIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSIATFFTGFIVGFTRGWKLTL
                     VILAISPVLGLSAAIWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFGGQKKEL
                     ERYNKNLEEAKGIGIKKAITANISIGAAFLLIYASYALAFWYGTSLVLSSEYSIGQVL
                     TVFFSVLIGAFSIGQASPSIEAFANARGAAYEIFKIIDNKPSIDSYSKSGHKPDNIKG
                     NLEFKNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQRLYDPTD
                     GMVCIDGQDIRTINVRHLREITGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKE
                     ANAYDFIMKLPNKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTE
                     SEAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELMKEKGIY
                     FKLVTMQTRGNEIELENATGESKSESDALEMSPKDSGSSLIKRRSTRRSIHAPQGQDR
                     KLGTKEDLNENVPPVSFWRILKLNSTEWPYFVVGIFCAIINGGLQPAFSIIFSRIIGI
                     FTRDEDPETKRQNSNMFSVLFLVLGIISFITFFLQGFTFGKAGEILTKRLRYMVFRSM
                     LRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAVITQNIANLGTGIIISLIYG
                     WQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRTVVSLTR
                     EQKFEYMYAQSLQVPYRNSLRKAHIFGVSFSITQAMMYFSYAGCFRFGAYLVANEFMN
                     FQDVLLVFSAIVFGAMAVGQVSSFAPDYAKAKVSAAHVIMIIEKSPLIDSYSPHGLKP
                     NTLEGNVTFNEVVFNYPTRPDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERF
                     YDPLAGSVLIDGKEIKHLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSHEE
                     IMQAAKEANIHHFIETLPEKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEA
                     TSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGKVKEHGTHQQL
                     LAQKGIYFSMVSVQAGAKR"
BASE COUNT     1296 a    833 c   1008 g   1142 t
ORIGIN

Alignment Scores:
Pred. No.:               1.55e-117       Length:          4279
Score:                   1557.50         Matches:         303
Percent Similarity:      72.12%          Conservative:    98
Best Local Similarity:   54.50%          Mismatches:      104
Query Match:             56.86%          Indels:          51
DB:                      6               Gaps:            5

US-09-873-409-7 (1-541) x AX105080 (1-4279)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            |||::::::|||  . :::|||::::::::::::|||||   ||||||||||||   |||
Db      167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
            |||   |||   |||:::|||:::         :::        :::
Db      227 GTTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy       35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                                 :::         ||||||      |||
Db      287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy       45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                                      ||||||   |||||||||         |||:::
Db      347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy       60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
                ||||||||:::|||   |||   :::   |||   ||||||      :::||||||
Db      407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy       80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
            ||||||||||||||||:::::::::      |||:::|||||||||||||||   |||:::
Db      467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy      100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
            ||||||||||||||||||:::|||   |||:::   ||||||::::::||||||||||||
Db      527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586
```

```
Qy     119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db     587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy     139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db     647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy     159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db     707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy     179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db     767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy     199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
Db     827 GGACAAAAGAAAGAACTTGAA------------------------------------- 847

Qy     219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
Db     848 ------------------------------------------AGGTACAACAAAAATTTA 865

Qy     239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
Db     866 GAAGAAGCTAAAGGAATTGGGATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCC 925

Qy     259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
Db     926 GCTTTCTTATTGATCTATGCATCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTC 985

Qy     279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
Db     986 CTCTCCAGTGAA------TATTCTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTA 1039

Qy     299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
Db    1040 ATTGGGGCTTTTAGTATTGGACAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGA 1099

Qy     319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
Db    1100 GGAGCAGCTTATGAAATCTTCAAGATAATTGACAATAAACCAAGCATTGACAGCTATTCG 1159

Qy     339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
Db    1160 AAGAGTGGACATAAACCAGATAATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTC 1219

Qy     359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
Db    1220 AGTTACCCTTCTCGAAAAGAAGTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGT 1279

Qy     379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
Db    1280 GGGCAGACAGTGGCGCTGGTTGGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTG 1339

Qy     399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
Db    1340 ATGCAGAGGCTCTATGACCCCACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGG 1399

Qy     419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db    1400 ACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTG 1459

Qy     439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
Db    1460 TTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAG 1519

Qy     459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
Db    1520 ATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTTATCATGAAACTACCTAATAAA 1579

Qy     479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db    1580 TTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATC 1639

Qy     499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db    1640 GCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCA 1699

Qy     519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db    1700 GCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCCCTGGATAAG 1747

RESULT 12
AX108656
LOCUS       AX108656                4195 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 3 from Patent WO0123565.
ACCESSION   AX108656
VERSION     AX108656.1  GI:13923888
KEYWORDS    .
SOURCE      crab-eating macaque.
  ORGANISM  Macaca fascicularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 4195)
  AUTHORS   Stocker,P.J., Steimel-Crespi,D.T. and Crespi,C.L.
  TITLE     P-glycoproteins from macaca fascicularis and uses thereof
  JOURNAL   Patent: WO 0123565-A 3 05-APR-2001;
            GENTEST CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..4195
                     /organism="Macaca fascicularis"
                     /db_xref="taxon:9541"
     CDS             100..3951
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC37765.1"
                     /db_xref="GI:13923889"
                     /translation="MDLEGDRNGGAEKKNFFKLNNKSKKDKKERKPTVSVFSMFRYSN
                     WLDKLYMVVGTLAAIIHGAGLPLMMLVFGDMTDTFANAGNLGDLGALLFNNTNSSNIT
                     DTVPVMNLEEDMTRYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQIHKIRKQFFHAIMR
                     QEIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSMATFFTGFIVGFTRGWKLT
                     LVILAISPVLGLSAAVWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFGGQKKE
                     LERYNKNLEEAKRIGIKKAITANISIGAAFLLIYASYALAFWYGTTLVLSKEYSIGQV
                     LTVFFSVLIGAFSVGQASPSIEAFANARGAAFEIFKIIDNKPSIDSYSKSGHKPDNIK
                     GNLEFRNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQRLYDPT
                     EGMVSVDGQDIRTINVRFLREIIGVVSQEPVLFATTIAENIRYGREDVTMDEIEKAVK
                     EANAYDFIMKLPQKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDT
                     ESEAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELMKEKGI
                     YFKLVTMQTAGNEIELENAADESKSEIDTLEMSSHDSGSSLIRKRSTRRSVRGSQGQD
                     RKLSTKEALDESIPPVSFWRIMKLNLTEWPYFVVGVFCAIINGGLQPAFAVIFSKIIG
                     IFTRNDDAETKRQNSNLFSLLFLVLGIVSFITFFLQGFTFGKAGEILTKRLRYMVFRS
                     MLRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAIITQNIANLGTGIIISLIY
                     GWQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRTVVSLT
                     QEQKFEHMYDQSLQVPYRNSLRKAHIFGITFSFTQAMMYFSYAGCFRFGAYLVAHSLM
                     SFEDVLLVFSAVVFGAMAVGQVSSFAPDYAKAKVSAAHIIMIIEKTPLIDSYSTEGLK
                     PNTLEGNVTFNEVVFNYPTRLDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLER
                     FYDPLAGKVLLDGKEIKQLNVQWLRAHLGIVSQEPILFDCSISENIAYGDNSRVVSQE
                     EIVRAAKEANIHAFIESLPNKYSTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDE
                     ATSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGRVKEHGTHQQ
                     LLAQKGIYFSMVSVQAGAKRQ"
BASE COUNT     1230 a    801 c   1039 g   1125 t
ORIGIN

Alignment Scores:
Pred. No.:              1.67e-117      Length:         4195
Score:                  1557.00        Matches:        306
Percent Similarity:     72.17%         Conservative:   96
Best Local Similarity:  54.94%         Mismatches:     103
Query Match:            56.85%         Indels:         52
DB:                     6              Gaps:           5

US-09-873-409-7 (1-541) x AX108656 (1-4195)
```

```
Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    250 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGAGCTGGACTTCCTCTCATGATGCTG 309

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSer---------------------GlyCys 33
Db    310 GTGTTTGGAGACATGACGGATACCTTTGCAAATGCAGGAAATTTAGGAGATTTAGGAGCT 369

Qy     34 LeuValGlnThrAsnThrTyrSer------------------------------------ 41
Db    370 CTGTTGTTTAACAACACTAATAGCAGTAATATCACTGATACAGTGCCCGTCATGAATCTG 429

Qy     42 ---------PhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIle 58
Db    430 GAGGAAGATATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTT 489

Qy     59 PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIle 78
Db    490 GCTGCTTACATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATT 549

Qy     79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp 98
Db    550 AGAAAACAGTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGAT 609

Qy     99 IleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly 117
Db    610 GTTGGGGAGCTTAACACCCGGCTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGT 669

Qy    118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
Db    670 GACAAAATTGGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGA 729

Qy    138 LeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
Db    730 TTTACACGTGGTTGGAAGCTAACCCTTGTGATTTGGCCATCAGTCCTGTTCTTGGACTG 789

Qy    158 SerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr 177
Db    790 TCAGCTGCAGTCTGGGCAAAGATACTGTCTTCATTTACTGATAAAGAACTCTTAGCTTAT 849

Qy    178 SerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197
Db    850 GCAAAAGCTGGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTT 909

Qy    198 ArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrp 217
Db    910 GGAGGACAAAAGAAAGAACTCGAA---------------------------------- 933

Qy    218 PheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsn 237
Db    934 ---------------------------------------------AGGTACAACAAAAAT 948

Qy    238 LeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGly 257
Db    949 TTAGAAGAAGCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGT 1008

Qy    258 AlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeu 277
Db   1009 GCTGCTTTCCTGCTTATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTG 1068

Qy    278 IleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerVal 297
Db   1069 GTCCTCTCAAAGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTA 1122

Qy    298 IleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAla 317
Db   1123 TTAATTGGGGCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCA 1182

Qy    318 ArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPhe 337
Db   1183 AGAGGAGCAGCTTTTGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTAT 1242

Qy    338 SerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSer 357
Db   1243 TCGAAGAGTGGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCAC 1302

Qy    358 PheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLys 377
Db   1303 TTCAGTTACCCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAG 1362

Qy    378 SerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGln 397
Db   1363 AGTGGGCAGACGGTGGCCCTGGTTGGAAACAGCGGCTGTGGGAAGAGCACAACGGTCCAG 1422

Qy    398 LeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIle 417
Db   1423 CTGATGCAGAGGCTTTATGACCCCACAGAGGGCATGGTCAGTGTTGATGGACAGGATATT 1482

Qy    418 ArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProVal 437
Db   1483 AGGACCATAAACGTAAGGTTTCTACGGGAAATCATCGGTGTGGTGAGTCAGGAACCTGTA 1542

Qy    438 LeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGlu 457
Db   1543 TTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGTCGTGAAGATGTCACCATGGAT 1602

Qy    458 GluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsn 477
Db   1603 GAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCAG 1662

Qy    478 LysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArg 497
Db   1663 AAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGCTGAGTGGTGGGCAGAAGCAGAGG 1722

Qy    498 IleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThr 517
Db   1723 ATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGACGAGGCCACG 1782

Qy    518 SerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db   1783 TCAGCCTTGGACACAGAAAGTGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 1833
```

```
RESULT 13
AF269224
LOCUS       AF269224                4045 bp    DNA     linear   SYN 04-JUL-2000
DEFINITION  Synthetic construct Canis familiaris his-tagged-multidrug
            resistance glycoprotein gene, complete cds.
ACCESSION   AF269224
VERSION     AF269224.1  GI:8926216
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4045)
  AUTHORS   Roulet,A.C., Puel,O., Gesta,S., Alvinerie,M. and Pineau,T.J.
  TITLE     Analysis of dog MDR1 p-glycoprotein
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4045)
  AUTHORS   Roulet,A.C., Puel,O., Gesta,S., Alvinerie,M. and Pineau,T.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-MAY-2000) Pharmacology, INRA, 180 Chemin de
            Tournefeuille, BP3, Toulouse 31931, France
FEATURES             Location/Qualifiers
     source          1..4045
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="original mRNA isolated from Beagle dog colon"
     CDS             101..3964
                     /codon_start=1
                     /transl_table=11
                     /product="his-tagged-multidrug resistance glycoprotein
                     MDR1"
                     /protein_id="AAF81747.1"
                     /db_xref="GI:8926217"
                     /translation="MHHHHHHDPEGGRKGSAEKNFWKMGKKSKKNEKKEKKPTVSTFA
```

```
                    MFRYSNWLDRLYMLVGTMAAIIHGAALPLMMLVFGNMTDSFANAGISRNKTFPVIINE
                    SITNNTQHFINHLEEEMTTYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQILKIRKQFF
                    HAIMRQEIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSIATFFTGFIVGFTR
                    GWKLTLVILAISPVLGLSAAIWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFG
                    GQKKELERYNKNLEEAKRIGIKKAITANISIGAAFLLIYASYALAFWYGTSLVLSSEY
                    SIGQVLTVFFSVLIGAFSIGQASPSIEAFANARGAAYEIFKIIDNKPSIDSYSKSGHK
                    PDNIKGNLEFKNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQR
                    LYDPTDGMVCIDGQDIRTINVRHLREITGVVSQEPVLFATTIAENIRYGRENVTMDEI
                    EKAVKEANAYDFIMKLPNKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEAT
                    SALDTESEAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELM
                    KEKGIYFKLVTMQTRGNEIELENATGESKSESDALEMSPKDSGSSLIKRRSTRRSIHA
                    PQGQDRKLGTKEDLNENVPPVSFWRILKLNSTEWPYFVVGIFCAIINGGLQPAFSIIF
                    SRIIGIFTRDEDPETKRQNSNMFSVLFLVLGIISFITFFLQGFTFGKAGEILTKRLRY
                    MVFRSMLRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAVITQNIANLGTGII
                    ISLIYGWQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRT
                    VVSLTREQKFEYMYAQSLQVPYRNSLRKAHIFGVSFSITQAMMYFSYAGCFRFGAYLV
                    ANEFMNFQDVLLVFSAIVFGAMAVGQVSSFAPDYAKAKVSAAHVIMIIEKSPLIDSYS
                    PHGLKPNTLEGNVTFNEVVFNYPTRPDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVV
                    QLLERFYDPLAGSVLIDGKEIKHLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSR
                    VVSHEEIMQAAKEANIHHFIETLPEKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHI
                    LLLDEATSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGKVKEH
                    GTHQQLLAQKGIYFSMISVQAGAKR"
BASE COUNT     1171 a    850 c    986 g   1038 t
ORIGIN

Alignment Scores:
Pred. No.:              1.75e-117      Length:        4045
Score:                  1556.50        Matches:       303
Percent Similarity:     72.12%         Conservative:  98
Best Local Similarity:  54.50%         Mismatches:    104
Query Match:            56.83%         Indels:        51
DB:                     12             Gaps:          5

US-09-873-409-7 (1-541) x AF269224 (1-4045)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    269 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 328

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
Db    329 GTTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 388

Qy     35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
Db    389 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 448

Qy     45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db    449 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 508

Qy     60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db    509 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 568

Qy     80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db    569 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 628

Qy    100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db    629 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 688

Qy    119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db    689 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 748

Qy    139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db    749 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 808

Qy    159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db    809 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 868

Qy    179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db    869 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 928

Qy    199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
Db    929 GGACAAAAGAAAGAACTTGAA--------------------------------------- 949

Qy    219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
Db    950 ------------------------------------------AGGTACAACAAAAATTTA 967

Qy    239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
Db    968 GAAGAAGCTAAAAGAATTGGGATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCC 1027

Qy    259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
Db   1028 GCTTTCTTATTGATCTATGCATCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTC 1087

Qy    279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
Db   1088 CTCTCCAGTGAA------TATTCTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTA 1141

Qy    299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
Db   1142 ATTGGGGCTTTTAGTATTGGACAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGA 1201

Qy    319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
Db   1202 GGAGCAGCTTATGAAATCTTCAAGATAATTGACAATAAACCAAGCATTGACAGCTATTCG 1261

Qy    339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
Db   1262 AAGAGTGGACATAAACCAGATAATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTC 1321

Qy    359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
Db   1322 AGTTACCCTTCTCGAAAAGAAGTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGT 1381

Qy    379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
Db   1382 GGGCAGACAGTGGCGCTGGTTGGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTG 1441

Qy    399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
Db   1442 ATGCAGAGGCTCTATGACCCCACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGG 1501

Qy    419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db   1502 ACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTG 1561

Qy    439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
Db   1562 TTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAG 1621

Qy    459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
Db   1622 ATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTTATCATGAAACTACCTAATAAA 1681

Qy    479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db   1682 TTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATC 1741

Qy    499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db   1742 GCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCA 1801

Qy    519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db   1802 GCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCCCTGGATAAG 1849
```

```
RESULT 14
AX322787
LOCUS       AX322787                3860 bp    DNA     linear   PAT 07-JAN-2002
DEFINITION  Sequence 1 from Patent WO0192877.
ACCESSION   AX322787
VERSION     AX322787.1  GI:18093766
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Sorrentino,B. and Schuetz,J.
  TITLE     Method of identifying and/or isolating stem cells
  JOURNAL   Patent: WO 0192877-A 1 06-DEC-2001;
            ST. JUDE CHILDREN'S RESEARCH HOSPITAL (US)
FEATURES             Location/Qualifiers
     source          1..3860
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1135 a    746 c    957 g   1022 t
ORIGIN

Alignment Scores:
Pred. No.:              1.99e-117      Length:         3860
Score:                  1555.50        Matches:        306
Percent Similarity:     73.29%         Conservative:   100
Best Local Similarity:  55.23%         Mismatches:     99
Query Match:            56.79%         Indels:         49
DB:                     6              Gaps:           6

US-09-873-409-7 (1-541) x AX322787 (1-3860)

Qy          1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
              |||::::::|||    |||||:::::::::::::||||||    |||||||||||     |||
Db        151 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210

Qy         21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
              |||   ||||||||:::|||                           :::|||:::|||
Db        211 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 270

Qy         32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                          ::::::      |||          ||||||
Db        271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330

Qy         44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                    |||              ||||||    ||||||||    |||:::    |||||
Db        331 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 390

Qy         62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
              ||||||:::|||    |||    :::   |||   ||||||       :::||||||||||
Db        391 ATTCAGGTTTCATTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450

Qy         82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
              |||||||||:::::::::    |||:::||||||||||||||||     |||:::||||||
Db        451 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 510

Qy        102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
              ||||||||||||:::|||   |||:::   ||||||::::::||||||||||||||||||
Db        511 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGGTATTGGTGACAAAATT 570

Qy        121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
              :::         ||||||:::|||:::||||||      |||        ||||||    :::
Db        571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 630

Qy        141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
              |||||||||||||||||||||     |||:::   ||||||::::::      |||||||||
Db        631 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

Qy        161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
                    ::::::::::::   |||   |||   |||||||||   ||||||:::||||||
Db        691 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy        181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
              ||||||||||||||||||||||||::::::||||||||||||||||||||        |||
Db        751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810

Qy        201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
              :::|||||||||:::
Db        811 AAGAAAGAACTTGAA--------------------------------------------- 825

Qy        221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                                 ||||||   :::||||||::::::
Db        826 -----------------------------------AGGTACAACAAAAATTTAGAAGAA 849

Qy        241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
              ||||||    ||||||||:::   |||   :::   :::|||:::||||||       :::
Db        850 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 909

Qy        261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
                 :::        :::|||  |||||||||||||||||||:::|||:::|||:::
Db        910 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 969

Qy        281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
              ||||||      |||:::||||||    ||||||    ||||||||||||||:::
Db        970 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1023

Qy        301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
              ::::::   :::|||   |||   |||        |||   ||||||   |||||||||||
Db       1024 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1083

Qy        321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
              |||:::    ||||||::::::||||||    ||||||||||||||::::::|||    :::
Db       1084 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1143

Qy        341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
              |||:::||||||::::::||||||:::|||:::||||||       |||:::|||
Db       1144 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1203

Qy        361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
              |||||||||      :::|||||||||||||||||||||||:::::::::||||||:::
Db       1204 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1263

Qy        381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
              ||||||||||||||||||     :::|||    ||||||||||||   ||||||:::|||
Db       1264 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1323

Qy        401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
              |||||||||||||||     :::|||    :::   ||||||      |||||||||    :::
Db       1324 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1383

Qy        421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
              |||||||||      |||:::   ||||||||||||||||||||||||||||||
Db       1384 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1443

Qy        441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
              ||||||||:::   ||||||:::||||||||::::::||||||   :::|||:::|||
Db       1444 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1503

Qy        461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
              :::|||   :::||||||||||||||||||||||||||||:::   |||:::||||||:::
Db       1504 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1563

Qy        481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
              |||||||||||||||:::|||||||||:::||||||||||||||||||||||||||||
Db       1564 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 1623

Qy        501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
              |||||||||||||||||||||||:::|||||||||||||||||||||||||||||||||
Db       1624 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 1683

Qy        521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
              |||::::||||||::::::   ||||||   ||||||:::|||
Db       1684 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 1725
```

```
RESULT 15
E02326
LOCUS       E02326                  4378 bp    RNA     linear   PAT 29-SEP-1997
DEFINITION  Multidrug resistance relating gene derived from human normal cells.
ACCESSION   E02326
VERSION     E02326.1  GI:2170561
KEYWORDS    JP 1990100680-A/1.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4378)
  AUTHORS   Ueda,K. and Komano,T.
  TITLE     HUMAN NORMAL CELL-DERIVED MDR RELATED GENE
  JOURNAL   Patent: JP 1990100680-A 1 12-APR-1990;
            SUNTORY LTD
COMMENT     OS   Homo sapiens
            PN   JP 1990100680-A/1
            PD   12-APR-1990
            PF   05-OCT-1988 JP 1988251475
            PI   UEDA KAZUMITSU, KOMANO TORU
            PC   C12N15/12,C12N1/21,C12Q1/68;
            CC   strandedness: Single;
            CC   topology: Linear;
            CC   *source: tissue_type=Adrenal gland;
            CC   *source: clone=SAM1132;
            FH   Key             Location/Qualifiers
            FH
            FT   5'UTR           1. .137
            FT   CDS             138. .3980
            FT                   /gene='Multidrug resistance relating gene' FT
             mat_peptide     138. .3977
            FT                   /gene='Multidrug resistance relating gene' FT
                             /gene='MDRA1'
            FT   3'UTR           3981. .>4378.
FEATURES             Location/Qualifiers
     source          1. .4378
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1315 a    818 c   1057 g   1188 t
ORIGIN

Alignment Scores:
Pred. No.:              2.33e-117      Length:         4378
Score:                  1555.50        Matches:        306
Percent Similarity:     73.29%         Conservative:   100
Best Local Similarity:  55.23%         Mismatches:     99
Query Match:            56.79%         Indels:         49
DB:                     6              Gaps:           6

US-09-873-409-7 (1-541) x E02326 (1-4378)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    288 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 347

Qy     21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db    348 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 407

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db    408 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 467

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    468 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 527

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    528 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 587

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    588 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 647

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    648 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAAATT 707

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    708 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 767

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    768 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 827

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    828 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 887

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    888 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 947

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db    948 AAGAAAGAACTTGAA--------------------------------------------- 962

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db    963 -----------------------------------AGGTACAACAAAAATTTAGAAGAA 986

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db    987 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 1046

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db   1047 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 1106

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db   1107 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1160

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db   1161 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1220

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   1221 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1280

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   1281 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1340

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   1341 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGTCTGAACCTGAAGGTGCAGAGTGGGCAG 1400

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   1401 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1460

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   1461 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1520

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   1521 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1580

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
```

```
Db   1581 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1640

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
          :::|||   :::|||||||||||||||||||||||||||:::   |||:::||||||:::
Db   1641 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1700

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
          |||||||||||||||:::|||||||||:::||||||||||||||||||||||||||||||
Db   1701 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 1760

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
          |||||||||||||||||||||||||||||||||:::||||||||||||||||||||||||
Db   1761 GCACGTGCCCTGGTTCGTAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 1820

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
          |||:::||||||::::::   ||||||   ||||||:::|||
Db   1821 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 1862

Search completed: March 31, 2003, 02:23:23
Job time : 7369.73 secs
```

GenCore version 5.1.4_p5_4578


```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 30, 2003, 01:20:01 ; Search time 528.737 Seconds
                                          (without alignments)
                                          2304.231 Million cell updates/sec

Title:          US-09-873-409-7
Perfect score:  2739
Sequence:       1 MILGILASLVNGACLPLMPL..........SESKSAVQAALEKDTPRYSF 541

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters:      4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115416_19216/app_query.fasta_1.7544
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_4247_@runat_27032003_115416_19216 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      N_Geneseq_101002:*
                1:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
                2:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
                3:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
                4:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
                5:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
                6:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
                7:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
                8:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
                9:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
                10:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
                11:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
                12:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
                13:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
                14:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
                15:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
                16:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
                17:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
                18:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
                19:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
                20:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
                21:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
                22:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
                23:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
                24:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2437	89.0	3699	24	ABK83223	Human transporter
2	1566.5	57.2	4189	21	AAZ49334	Murine multidrug r
3	1566.5	57.2	4189	24	ABA94367	Mouse BCRP DNA rel
4	1566.5	57.2	4313	14	AAQ38950	Mouse multidrug re
5	1561.5	57.0	4279	22	AAD03506	Dog P-glycoprotein
6	1557.5	56.9	4279	22	AAD03504	Dog P-glycoprotein
7	1557.5	56.9	4279	22	AAD03505	Dog P-glycoprotein
8	1557	56.8	4195	22	AAF86128	Cynomologous monke
9	1555.5	56.8	3860	21	AAZ49332	Human wild-type mu
10	1555.5	56.8	3860	24	ABA94365	Human BCRP DNA rel
11	1552.5	56.7	4233	21	AAZ90198	Rat mdr1b2 (multis
12	1552.5	56.7	4233	22	AAF27498	Rat mdr1b2 multidr
13	1552.5	56.7	4279	22	AAD03488	Dog P-glycoprotein
14	1552.5	56.7	4317	22	AAD03489	Dog P-glycoprotein
15	1552.5	56.7	4669	19	AAV32645	Human P glycoprote
16	1552.5	56.7	4669	24	ABK52041	cDNA encoding huma
17	1548.5	56.5	4186	22	AAF86127	Cynomologous monke
18	1548	56.5	4254	24	ABK63517	Rat sequence diffe
19	1546.5	56.5	3860	21	AAZ49333	Human G185V mutant
20	1546.5	56.5	3860	24	ABA94366	Human BCRP DNA rel
21	1546.5	56.5	4349	22	AAH57442	Human intestine ce
22	1546.5	56.5	4646	15	AAQ72872	Human multidrug re
23	1546.5	56.5	4646	21	AAZ94738	Human ATP binding
24	1546.5	56.5	4646	24	AAD38994	Human mdr1 gene.
25	1546.5	56.5	4646	24	ABL68592	Kidney cancer rela
26	1546.5	56.5	4646	24	ABL68880	Kidney cancer rela
27	1546.5	56.5	6505	17	AAT13394	Hybrid vector pSF-
28	1546.5	56.5	8630	21	AAZ24041	Retroviral M4 mdr-
29	1546.5	56.5	8630	21	AAZ24042	Retroviral vector
30	1545	56.4	4264	19	AAV66533	Mutated human P-gl
31	1545	56.4	4264	19	AAV66534	Mutated human P-gl
32	1544.5	56.4	4669	8	AAN70752	Sequence of human
33	1543.5	56.4	4378	11	AAQ04522	Multidrug Resistan
34	1543	56.3	4425	21	AAZ52048	Rat multidrug resi
35	1541	56.3	4369	21	AAZ52047	Rat multidrug resi
36	1540.5	56.2	4669	14	AAQ52726	Sequence of human
37	1539	56.2	3840	24	ABL91687	Human polynucleoti
38	1539	56.2	3988	21	AAZ88973	Human MDR-1 DNA.
39	1536	56.1	4788	21	AAZ49335	Murine multidrug r
40	1536	56.1	4788	24	ABA94368	Mouse BCRP DNA rel
41	1455.5	53.1	3924	21	AAZ94742	Human ATP binding
42	1455.5	53.1	3924	21	AAZ88974	Human MDR-3 DNA.
43	1455.5	53.1	3924	24	ABN95801	Gene #2299 used to
44	1421.5	51.9	3912	24	ABK63653	Rat sequence diffe
45	1250	45.6	4776	21	AAZ94744	Human ATP binding

ALIGNMENTS

```
RESULT 1
ABK83223
ID   ABK83223 standard; cDNA; 3699 BP.
XX
AC   ABK83223;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human transporter and ion channel, TRICH14, Incyte ID 7472030CB1, cDNA.
XX
KW   Human; ss; gene; transporter and ion channel; TRICH; transport disorder;
KW   neurological disorder; muscle disorder; immunological disorder; cancer;
KW   scleroderma; systemic lupus erythematosus; allergy; leukaemia;
KW   cell proliferative disorder; cervical cancer; breast cancer;
KW   neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;
KW   myotonic dystrophy; catatonia; endocrine disorder; diabetes;
KW   Grave's disease; gastrointestinal disorder; Crohn's disease;
KW   renal disorder; Good pasture's syndrome; viral infection; cirrhosis;
```

```
KW   bacterial infection; fungal infection; parasitic infection;
KW   protozoal infection; helminthic infection; cardiovascular disorder;
KW   atherosclerosis; hepatic disease.
XX
OS   Homo sapiens.
XX
PN   WO200240541-A2.
XX
PD   23-MAY-2002.
XX
PF   25-OCT-2001; 2001WO-US46055.
XX
PR   27-OCT-2000; 2000US-243989P.
PR   03-NOV-2000; 2000US-245904P.
PR   09-NOV-2000; 2000US-247673P.
PR   17-NOV-2000; 2000US-249661P.
PR   20-NOV-2000; 2000US-252232P.
PR   01-DEC-2000; 2000US-250790P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Tang YT,  Yue H,  Nguyen DB,  Hafalia AJA,  Elliott VS,  Lu Y;
PI   Walia NK,  Yao MG,  Baughn MR,  Gandhi AR,  Ding L,  Sanjanwala M;
PI   Ramkumar J,  Arvizu C,  Gietzen KJ,  Lal PG,  Azimzai Y,  Khan FA;
PI   Thangavelu K,  Thornton M,  Lu DAM,  Tribouley CM,  Warren BA;
PI   Ison CH,  Das D,  Raumann BE,  Policky JL,  Kearney L;
XX
DR   WPI; 2002-463570/49.
DR   P-PSDB; ABG61544.
XX
PT   New transporters and ion channels (TRICH) polypeptides, useful for
PT   diagnosing, preventing, and treating disorders associated with an
PT   abnormal expression or activity of TRICH, e.g. immunological, muscular
PT   or renal disorders  -
XX
PS   Claim 5; Page 172; 178pp; English.
XX
CC   The invention relates to human transporters and ion channels (TRICH)
CC   polypeptides, a naturally occurring amino acid sequence 90 % identical to
CC   TRICH, a biologically active fragment of TRICH or an immunogenic fragment
CC   of TRICH. Also included are an isolated polynucleotide encoding TRICH,
CC   a recombinant polynucleotide comprising a promoter sequence operably
CC   linked to the TRICH polynucleotide, a cell transformed with the
CC   recombinant polynucleotide, a transgenic organism comprising the
CC   recombinant polynucleotide, an isolated antibody that binds specifically
CC   to TRICH, and screening for compounds which bind to TRICH, modulate
CC   TRICH, modulate TRICH expression or are ant/agonists of TRICH.
CC   The polypeptides are useful for diagnosing, treating, and
CC   preventing transport, neurological, muscle, immunological disorders
CC   (e.g. scleroderma, systemic lupus erythematosus, allergies), cell
CC   proliferative disorders such as cancers (e.g. leukaemia, cervical or
CC   breast cancers), neurodegenerative disorders (e.g. Parkinson's disease,
CC   Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy,
CC   catatonia), endocrine disorders (e.g. diabetes, Grave's disease),
CC   gastrointestinal disorders (e.g. Crohn's disease), renal disorders
CC   (e.g. Good pasture's syndrome), viral, bacterial, fungal, parasitic,
CC   protozoal and helminthic infections, cardiovascular disorders (e.g.
CC   atherosclerosis), or hepatic diseases (e.g. cirrhosis) and many
CC   other diseases and disorders detailed in the specification. They can also
CC   be used in assessing the effects of exogenous compounds on the
CC   expression of nucleic acid and amino acid sequences of transporters and
CC   ion channels. TRICH or its fragments may also be used in screening for
CC   compounds that specifically bind to and modulate the activity of TRICH.
CC   The polynucleotides can be used to create knock-in humanised animals or
CC   transgenic animals to model human disease. The present sequence
CC   encodes a TRICH protein.
XX
SQ   Sequence 3699 BP; 1116 A; 707 C; 860 G; 1016 T; 0 other;

Alignment Scores:
Pred. No.:            1.45e-241     Length:        3699
Score:                2437.00       Matches:       495
Percent Similarity:   90.71%        Conservative:  3
```

```
Best Local Similarity:  90.16%        Mismatches:   9
Query Match:            88.97%        Indels:       42
DB:                     24            Gaps:         3

US-09-873-409-7 (1-541) x ABK83223 (1-3699)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           ||||||||||||||||    ||||||   ||||||||||||||||||||||||||||||
Db     145 ATGATCCTGGGTATACTGACATCACTGTTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 204

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
              :::||||||||||||||||||||||||||||||||||||||    |||||||||
Db     205 TGTATAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCACACTAACACAACA 264

Qy      41 SerPhePheArg------------------------------------------LeuThr 46
           ::::::                                                ||||||
Db     265 AATTATCAGAACTGTACTCAGTCTCAAGAGAAGCTGAATGAAGATATGACTCTGTTGACC 324

Qy      47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     325 CTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGTTACATACAGATTTCCTTG 384

Qy      67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     385 TGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAACAGTTTTTTCATTCAGTT 444

Qy      87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     445 TTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGTGAACTTAACACTCGCATG 504

Qy     107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
           |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     505 ACAGATGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATTGCTCTGTTGTTTCAA 564

Qy     126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     565 AACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAGGGCTGGAAACTCACC 624

Qy     146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     625 CTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCAGCATGTTCTAGGATG 684

Qy     166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     685 GTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCAGAA 744

Qy     186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     745 GAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAGGAGAAAGAACTTCAA 804

Qy     206 ArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeu 225

Db     804 ------------------------------------------------------------ 804

Qy     226 SerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 245
                                 ||||||||||||||||||||||||||||||||||||||||
Db     805 ---------------------AGGTATACACAGAATCTCAAAGATGCAAAGGATTTTGGC 843

Qy     246 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 265
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     844 ATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAATGGA 903

Qy     266 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 285
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     904 ACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGATAT 963

Qy     286 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 305
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     964 ACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGTAGTTATTGCATTGGA 1023

Qy     306 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 325
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1024 GCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTC 1083
```

```
Qy     326 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 345
           |||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db    1084 CAGGTTATTGATAAGAAACCCAGTATAGGTAACTTTTCCACAGCTGGATATAAACCTGAA 1143

Qy     346 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 365
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1144 TCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGGCCATCT 1203

Qy     366 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 385
           |||||||||||||||||||||||||||   ||||||||||||||||||||||||||||||
Db    1204 ATCAAGATTCTGAAAGGTCTGAATCTCGGAATTAAGTCTGGAGAGACAGTCGCCTTGGTC 1263

Qy     386 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 405
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1264 GGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCG 1323

Qy     406 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 425
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1324 GATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTAT 1383

Qy     426 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 445
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1384 CGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAAC 1443

Qy     446 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 465
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1444 AATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAA 1503

Qy     466 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 485
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1504 GCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAA 1563

Qy     486 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 505
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1564 AAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTT 1623

Qy     506 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 525
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1624 CGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAG 1683

Qy     526 SerAlaValGlnAlaAlaLeuGluLys 534
           |||||||||||||||||||||||||||
Db    1684 TCAGCTGTTCAAGCTGCACTGGAGAAG 1710
```

RESULT 2
AAZ49334

```
ID   AAZ49334 standard; cDNA; 4189 BP.
XX
AC   AAZ49334;
XX
DT   14-MAR-2000  (first entry)
XX
DE   Murine multidrug resistance-1 (MDR-1) cDNA.
XX
KW   Multidrug resistance; MDR-1; P-glycoprotein;
KW   transmembrane efflux pump; haematopoietic stem cell; transduction;
KW   bone marrow transplantation; chemotherapy; radiation therapy; cancer;
KW   gene therapy; gene replacement; genetic defect; thalassaemia;
KW   Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
KW   cytokine; ds.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3831
FT                   /*tag=  a
FT                   /product= "Murine MDR-1 protein"
XX
PN   WO9961589-A2.
XX
PD   02-DEC-1999.
XX
PF   27-MAY-1999;   99WO-US11825.
XX
PR   28-MAY-1998;   98US-0086988.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI   Sorrentino B,  Bunting K;
XX
DR   WPI; 2000-072615/06.
DR   P-PSDB; AAY58188.
XX
PT   Ex vivo expansion of hematopoietic stem cells transduced with a
PT   sequence encoding human multidrug resistance-1, used for bone marrow
PT   transplantation -
XX
PS   Disclosure; Page 90-93; 113pp; English.
XX
CC   This sequence represents cDNA encoding murine multidrug
CC   resistance protein MDR-1. MDR-1 is a transmembrane
CC   efflux pump, responsible for the export of drugs from cells,
CC   particularly cancer cells. The invention relates to transducing
CC   haematopoietic stem cells with nucleic acid encoding an MDR protein
CC   and culturing the modified cells. The modified haematopoietic stem
CC   cells are useful in bone marrow transplantation (to reconstitute
CC   haematopoietic systems in patients who have undergone chemotherapy or
CC   radiation therapy) and in ex vivo gene therapy of genetic defects in
CC   cells derived from haematopoietic stem cells, e.g., thalassaemia,
CC   Gauchier's disease, sickle cell anaemia or leukaemia. The modified
CC   cells can also be used to identify factors involved in regulating
CC   proliferation and differentiation in haematopoietic stem cells.
CC   Haematopoietic stem cells that express MDR-1 will be protected against
CC   chemotherapeutic agents, so can be engrafted while the patient is
CC   undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells
CC   provides sufficient cells to permit standard biochemical analysis.
CC   Overexpression of MDR-1 allows cytokine-driven expansion of
CC   haematopoietic stem cells by at least 10-fold compared with a maximum
CC   of 4-fold in known procedures.
XX
SQ   Sequence 4189 BP; 1204 A; 875 C; 1028 G; 1082 T; 0 other;

Alignment Scores:
Pred. No.:              2.06e-151      Length:        4189
Score:                  1566.50        Matches:       307
Percent Similarity:     72.63%         Conservative:  99
Best Local Similarity:  54.92%         Mismatches:    94
Query Match:            57.19%         Indels:        59
DB:                     21             Gaps:          6

US-09-873-409-7 (1-541) x AAZ49334 (1-4189)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||||||||||    |||||:::::::::::|||         ||||||||:::    |||
Db     148 ATGATTCTGGGAACTCTCGCTGCTATTATCCATGGAACATTACTTCCCCTCTTGATGCTG 207

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
           |||   |||   |||:::|||:::
Db     208 GTGTTTGGAAACATGACAGATAGTTTACAAAAGCAGAAGCCAGTATTCTGCCAAGCATT 267

Qy      31 -----------------SerGlyCysLeuValGlnThrAsnThr--------------- 39
                            :::       |||:::   :::|||:::
Db     268 ACTAATCAAAGTGGACCCAACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAAGAG 327

Qy      40 ---------TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56
                    |||::::::              |||||   ||||||||
Db     328 ATGGCCATATACGCCTAC---------------TATTACACCGGGATTGGTGCTGGTGTG 372

Qy      57 LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys 76
           ||||||    |||||||:::||||||||   :::   |||   ||||||
Db     373 CTCATAGTTGCCTACATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGAAGACAGATACAC 432

Qy      77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96
           :::||||||::::::||||||||:::::::::   |||:::||||||||||||||
Db     433 AAGATTAGGCAGAAGTTTTTCCATGCTATAATGAATCAGGAGATAGGCTGGTTTGATGTG 492
```

```
Qy     97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
Db    493 CATGATGTTGGGGAGCTCAACACCCGGCTCACAGATGATGTCTCCAAAATTAATGACGGA 552

Qy    116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
Db    553 ATTGGTGACAAAATTGGGATGTTTTTCAGTCCATAACCACATTTTAGCCGGTTTTATC 612

Qy    136 ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
Db    613 ATAGGATTTATAAGTGGTTGGAAGCTAACCCTTGTCATTTTGGCTGTCAGCCCTCTTATT 672

Qy    156 MetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
Db    673 GGATTGTCATCTGCTTTGTGGGCAAAGGTATTGACTTCATTTACTAATAAGGAACTCCAG 732

Qy    176 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIle 195
Db    733 GCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT 792

Qy    196 AlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyr 215
Db    793 GCCTTTGGAGGACAACAGAAGGAACTTGAA------------------------------ 822

Qy    216 AlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThr 235
Db    823 ---------------------------------------------------AGGTACAAT 831

Qy    236 GlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSer 255
Db    832 AAAAATTTAGAAGAAGCTAAAAATGTTGGCATAAAGAAAGCTATCACAGCCAGCATTTCG 891

Qy    256 LeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThr 275
Db    892 ATAGGCATTGCCTACCTGTTGGTCTATGCATCATATGCACTGGCATTCTGGTATGGGACA 951

Qy    276 SerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePhe 295
Db    952 TCCTTGGTCCTCTCAAATGAA------TATTCTATTGGAGAAGTGCTTACTGTCTTCTTC 1005

Qy    296 SerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAla 315
Db   1006 TCTATTTTGTTGGGGACTTTTAGTATTGGACACTTGGCCCCAAACATAGAAGCCTTTGCA 1065

Qy    316 IleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAsp 335
Db   1066 AACGCACGAGGGGCAGCCTTTGAAATCTTCAAGATAATTGATAACGAGCCAAGCATTGAC 1125

Qy    336 AsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsn 355
Db   1126 AGCTTCTCAACAAAGGGCTACAAACCAGACAGTATAATGGGAAACTTAGAGTTTAAAAAT 1185

Qy    356 ValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArg 375
Db   1186 GTTCACTTCAACTACCCATCGAGAAGCGAAGTTCAGATCTTGAAGGGCCTCAATCTGAAG 1245

Qy    376 IleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVal 395
Db   1246 GTGAAGAGCGGACAGACGGTGGCCTTGGTTGGCAACAGTGGCTGTGGAAAAAGCACAACT 1305

Qy    396 ValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsn 415
Db   1306 GTCCAGCTGATGCAGAGGCTCTACGACCCCCTGGAGGGCGTGGTCAGTATCGACGGACAA 1365

Qy    416 AspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGlu 435
Db   1366 GACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCATTGGTGTGGTGAGTCAGGAA 1425

Qy    436 ProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThr 455
Db   1426 CCTGTGCTGTTTGCCACCACGATCGCCGAGAACATTCGCTATGGCCGAGAAGATGTCACC 1485

Qy    456 AspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhe 475
Db   1486 ATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTCATCATGAAACTG 1545

Qy    476 ProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLys 495
Db   1546 CCCCACCAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAA 1605

Qy    496 GlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGlu 515
Db   1606 CAGAGAATCGCCATTGCCCGGGCCCTGGTCCGCAATCCCAAGATCCTTTTGTTGGACGAG 1665

Qy    516 AlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db   1666 GCCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTGGTGCAGGCCGCACTGGATAAG 1722
```

```
RESULT 3
ABA94367
ID   ABA94367 standard; DNA; 4189 BP.
XX
AC   ABA94367;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Mouse BCRP DNA related seq Id No. 5.
XX
KW   Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;
KW   hepatotropic; neurodegenerative; cytostatic; antianemic; muscular; BCRP;
KW   cardiant; gene therapy; ds.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3831
FT                   /*tag=  a
XX
PN   WO200192877-A2.
XX
PD   06-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US17459.
XX
PR   31-MAY-2000; 2000US-0584586.
PR   29-MAY-2001; 2001US-0866866.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI   Sorrentino B,  Schuetz J;
XX
DR   WPI; 2002-114368/15.
DR   P-PSDB; ABB07268.
XX
PT   Identifying a stem cell, for treating e.g., muscular dystrophy,
PT   myocardial infarction, Parkinson's disease, or neurodegenerative
PT   disorders, comprises detecting the expression of an ATP transport
PT   protein (BCRP) by a cell  -
XX
PS   Disclosure; Page 64-66; 87pp; English.
XX
CC   The invention provides a method of identifying and/or isolating a stem
CC   cell that involves detecting the expression of an ATP transport protein
CC   containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
CC   comprising stem cells. The isolated stem cells may be used in the
CC   treatment of diseases such as muscular dystrophy, degenerative liver
CC   disorder, myocardial infarction, Parkinson's disease, degenerative
CC   disorders of the brain, and for tissue regeneration or replacement.
CC   Haematopoietic cells can be used in bone marrow transplants (e.g., for
CC   treatment of leukemia) and for ex vivo gene therapy for treating blood
CC   diseases such as sickle cell anemia and thalassemia. The stem cells can
CC   also be used as cell targets in gene therapy protocols. The present
CC   sequence represents a sequence related to the BCRP for which no relevant
CC   information has been provided in the specification.
XX
```

```
SQ   Sequence 4189 BP; 1204 A; 875 C; 1028 G; 1082 T; 0 other;

Alignment Scores:
Pred. No.:              2.06e-151    Length:        4189
Score:                  1566.50      Matches:       307
Percent Similarity:     72.63%       Conservative:  99
Best Local Similarity:  54.92%       Mismatches:    94
Query Match:            57.19%       Indels:        59
DB:                     24           Gaps:          6

US-09-873-409-7 (1-541) x ABA94367 (1-4189)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    148 ATGATTCTGGGAACTCTCGCTGCTATTATCCATGGAACATTACTTCCCCTCTTGATGCTG 207

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
Db    208 GTGTTTGGAAACATGACAGATAGTTTTACAAAAGCAGAAGCCAGTATTCTGCCAAGCATT 267

Qy     31 ------------------SerGlyCysLeuValGlnThrAsnThr--------------- 39
Db    268 ACTAATCAAAGTGGACCCAACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAAGAG 327

Qy     40 ---------TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56
Db    328 ATGGCCATATACGCCTAC---------------TATTACACCGGGATTGGTGCTGGTGTG 372

Qy     57 LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys 76
Db    373 CTCATAGTTGCCTACATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGAAGACAGATACAC 432

Qy     77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96
Db    433 AAGATTAGGCAGAAGTTTTTCCATGCTATAATGAATCAGGAGATAGGCTGGTTTGATGTG 492

Qy     97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
Db    493 CATGATGTTGGGGAGCTCAACACCCGGCTCACAGATGATGTCTCCAAAATTAATGACGGA 552

Qy    116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
Db    553 ATTGGTGACAAAATTGGGATGTTTTTCAGTCCATAACCACATTTTTAGCCGGTTTTATC 612

Qy    136 ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
Db    613 ATAGGATTTATAAGTGGTTGGAAGCTAACCCTTGTCATTTTGGCTGTCAGCCCTCTTATT 672

Qy    156 MetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
Db    673 GGATTGTCATCTGCTTTGTGGGCAAAGGTATTGACTTCATTTACTAATAAGGAACTCCAG 732

Qy    176 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIle 195
Db    733 GCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT 792

Qy    196 AlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyr 215
Db    793 GCCTTTGGAGGACAACAGAAGGAACTTGAA------------------------------ 822

Qy    216 AlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThr 235
Db    823 ---------------------------------------------------AGGTACAAT 831

Qy    236 GlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSer 255
Db    832 AAAAATTTAGAAGAAGCTAAAAATGTTGGCATAAAGAAAGCTATCACAGCCAGCATTTCG 891

Qy    256 LeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThr 275
Db    892 ATAGGCATTGCCTACCTGTTGGTCTATGCATCATATGCACTGGCATTCTGGTATGGGACA 951

Qy    276 SerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePhe 295
Db    952 TCCTTGGTCCTCTCAAATGAA------TATTCTATTGGAGAAGTGCTTACTGTCTTCTTC 1005

Qy    296 SerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAla 315
Db   1006 TCTATTTTGTTGGGGACTTTTAGTATTGGACACTTGGCCCCAAACATAGAAGCCTTTGCA 1065

Qy    316 IleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAsp 335
Db   1066 AACGCACGAGGGGCAGCCTTTGAAATCTTCAAGATAATTGATAACGAGCCAAGCATTGAC 1125

Qy    336 AsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsn 355
Db   1126 AGCTTCTCAACAAAGGGCTACAAACCAGACAGTATAATGGGAAACTTAGAGTTTAAAAAT 1185

Qy    356 ValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArg 375
Db   1186 GTTCACTTCAACTACCCATCGAGAAGCGAAGTTCAGATCTTGAAGGGCCTCAATCTGAAG 1245

Qy    376 IleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVal 395
Db   1246 GTGAAGAGCGGACAGACGGTGGCCTTGGTTGGCAACAGTGGCTGTGGAAAAAGCACAACT 1305

Qy    396 ValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsn 415
Db   1306 GTCCAGCTGATGCAGAGGCTCTACGACCCCCTGGAGGGCGTGGTCAGTATCGACGGACAA 1365

Qy    416 AspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGlu 435
Db   1366 GACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCATTGGTGTGGTGAGTCAGGAA 1425

Qy    436 ProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThr 455
Db   1426 CCTGTGCTGTTTGCCACCACGATCGCCGAGAACATTCGCTATGGCCGAGAAGATGTCACC 1485

Qy    456 AspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhe 475
Db   1486 ATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTCATCATGAAACTG 1545

Qy    476 ProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLys 495
Db   1546 CCCCACCAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAA 1605

Qy    496 GlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGlu 515
Db   1606 CAGAGAATCGCCATTGCCCGGGCCCTGGTCCGCAATCCCAAGATCCTTTTGTTGGACGAG 1665

Qy    516 AlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db   1666 GCCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTGGTGCAGGCCGCACTGGATAAG 1722

RESULT 4
AAQ38950
ID   AAQ38950 standard; DNA; 4313 BP.
XX
AC   AAQ38950;
XX
DT   28-JUL-1993  (first entry)
XX
DE   Mouse multidrug resistance sequence.
XX
KW   mdr gene; Lambda DR11 clone; ss.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   CDS             110..3940
FT                   /*tag= a
FT                   /phenotype= multidrug_resistance
XX
PN   US5198344-A.
XX
PD   30-MAR-1993.
```

```
XX
PF   15-JUL-1986;   86US-0885951.
XX
PR   15-JUL-1986;   86US-0885951.
PR   06-FEB-1991;   91US-0652311.
XX
PA   (MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
PI   Croop JM,  Gros P,  Housman DE;
XX
DR   WPI; 1993-126077/15.
DR   P-PSDB; AAR35199.
XX
PT   DNA sequence which confers multi-drug resistance on sensitive
PT   mammalian cells - used to preserve bone marrow cells during
PT   chemotherapy to prevent infection
XX
PS   Claim 1; Fig 8; 22pp; English.
XX
CC   A cDNA library was constructed from mRNA which had been isolated
CC   from a drug-sensitive mouse cell-line. Two mouse cDNA molecules
CC   complementary to the mRNA species encoded by 2 related but distinct
CC   mdr genes were isolated and cloned. One of the cDNA clones (lambda
CC   DR11) is a full-length cDNA clone for one member of the mdr gene
CC   family. When incorporated into prokaryotic expression vector pDREX4
CC   (which allows high levels of transcription of the cDNA when
CC   introduced into mammalian cells), the clone was shown to confer the
CC   multidrug resistance phenotype upon transfection into drug-sensitive
CC   mammalian cells.
XX
SQ   Sequence 4313 BP; 1241 A; 904 C; 1058 G; 1110 T; 0 other;

Alignment Scores:
Pred. No.:              2.15e-151     Length:         4313
Score:                  1566.50       Matches:        307
Percent Similarity:     72.63%        Conservative:   99
Best Local Similarity:  54.92%        Mismatches:     94
Query Match:            57.19%        Indels:         59
DB:                     14            Gaps:           6

US-09-873-409-7 (1-541) x AAQ38950 (1-4313)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    257 ATGATTCTGGGAACTCTCGCTGCTATTATCCATGGAACATTACTTCCCCTCTTGATGCTG 316

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
Db    317 GTGTTTGGAAACATGACAGATAGTTTTACAAAAGCAGAAGCCAGTATTCTGCCAAGCATT 376

Qy     31 ------------------SerGlyCysLeuValGlnThrAsnThr--------------- 39
Db    377 ACTAATCAAAGTGGACCCAACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAAGAG 436

Qy     40 ---------TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56
Db    437 ATGGCCATATACGCCTAC---------------TATTACACCGGGATTGGTGCTGGTGTG 481

Qy     57 LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys 76
Db    482 CTCATAGTTGCCTACATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGAAGACAGATACAC 541

Qy     77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96
Db    542 AAGATTAGGCAGAAGTTTTTCCATGCTATAATGAATCAGGAGATAGGCTGGTTTGATGTG 601

Qy     97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
Db    602 CATGATGTTGGGGAGCTCAACACCCGGCTCACAGATGATGTCTCCAAAATTAATGACGGA 661

Qy    116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
Db    662 ATTGGTGACAAAATTGGGATGTTTTTTCAGTCCATAACCACATTTTTAGCCGGTTTTATC 721

Qy    136 ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
Db    722 ATAGGATTTATAAGTGGTTGGAAGCTAACCCTTGTCATTTTGGCTGTCAGCCCTCTTATT 781

Qy    156 MetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
Db    782 GGATTGTCATCTGCTTTGTGGGCAAAGGTATTGACTTCATTTACTAATAAGGAACTCCAG 841

Qy    176 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIle 195
Db    842 GCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT 901

Qy    196 AlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyr 215
Db    902 GCCTTTGGAGGACAACAGAAGGAACTTGAA------------------------------ 931

Qy    216 AlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThr 235
Db    932 ---------------------------------------------------AGGTACAAT 940

Qy    236 GlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSer 255
Db    941 AAAAATTTAGAAGAAGCTAAAAATGTTGGCATAAAGAAAGCTATCACAGCCAGCATTTCG 1000

Qy    256 LeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThr 275
Db   1001 ATAGGCATTGCCTACCTGTTGGTCTATGCATCATATGCACTGGCATTCTGGTATGGACA 1060

Qy    276 SerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePhe 295
Db   1061 TCCTTGGTCCTCTCAAATGAA------TATTCTATTGGAGAAGTGCTTACTGTCTTCTTC 1114

Qy    296 SerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAla 315
Db   1115 TCTATTTTGTTGGGGACTTTTAGTATTGGACACTTGGCCCCAAACATAGAAGCCTTTGCA 1174

Qy    316 IleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAsp 335
Db   1175 AACGCACGAGGGGCAGCCTTTGAAATCTTCAAGATAATTGATAACGAGCCAAGCATTGAC 1234

Qy    336 AsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsn 355
Db   1235 AGCTTCTCAACAAAGGGCTACAAACCAGACAGTATAATGGGAAACTTAGAGTTTAAAAAT 1294

Qy    356 ValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArg 375
Db   1295 GTTCACTTCAACTACCCATCGAGAAGCGAAGTTCAGATCTTGAAGGGCCTCAATCTGAAG 1354

Qy    376 IleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVal 395
Db   1355 GTGAAGAGCGGACAGACGGTGGCCTTGGTTGGCAACAGTGGCTGTGGAAAAAGCACAACT 1414

Qy    396 ValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsn 415
Db   1415 GTCCAGCTGATGCAGAGGCTCTACGACCCCCTGGAGGGCGTGGTCAGTATCGACGGACAA 1474

Qy    416 AspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGlu 435
Db   1475 GACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCATTGGTGTGGTGAGTCAGGAA 1534

Qy    436 ProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThr 455
Db   1535 CCTGTGCTGTTTGCCACCACGATCGCCGAGAACATTCGCTATGGCCGAGAAGATGTCACC 1594

Qy    456 AspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhe 475
Db   1595 ATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTCATCATGAAACTG 1654

Qy    476 ProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLys 495
Db   1655 CCCCACCAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAA 1714
```

```
Qy     496 GlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGlu 515
               |||||||||||||||||||||||||||||||||||||||||||||||||:::|||||||||
Db    1715 CAGAGAATCGCCATTGCCCGGGCCCTGGTCCGCAATCCCAAGATCCTTTTGTTGGACGAG 1774

Qy     516 AlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
               |||||||||||||||||:::||||||::::::   ||||||||||||||:::|||
Db    1775 GCCACCTCAGCCCTGGATACAGAAAGTGAAGCTGTGGTGCAGGCCGCACTGGATAAG 1831
```

RESULT 5
AAD03506
ID   AAD03506 standard; cDNA; 4279 BP.
XX
AC   AAD03506;
XX
DT   13-JUN-2001 (first entry)
XX
DE   Dog P-glycoprotein (PGP) allelic variant (Genotype D) cDNA.
XX
KW   Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KW   MDR1; drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
```
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag= a
FT                   /product= "Dog P-glycoprotein (PGP) allelic variant
FT                   (Genotype D) protein"
FT   allele          replace (91, T)
FT                   /*tag= b
FT   allele          replace (607, C)
FT                   /*tag= c
FT   allele          replace (1001, T)
FT                   /*tag= c
FT   allele          replace (3458, A)
FT                   /*tag= c
```
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ, Steimel-crespi DT, Crespi CL, Reif TC, Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00310.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 9; Page 102-107; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) allelic variant
CC   (Genotype D) cDNA. The PGP enzyme functions as an efflux pump
CC   exporting small molecules across the cell membrane. This enzyme
CC   is a member of the ABC transporter family.
XX
SQ   Sequence 4279 BP; 1296 A; 833 C; 1009 G; 1141 T; 0 other;

```
Alignment Scores:
Pred. No.:              6.99e-151      Length:         4279
Score:                  1561.50        Matches:        304
Percent Similarity:     72.12%         Conservative:   97
Best Local Similarity:  54.68%         Mismatches:     104
Query Match:            57.01%         Indels:         51
DB:                     22             Gaps:           5
```

US-09-873-409-7 (1-541) x AAD03506 (1-4279)

```
Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
               |||::::::|||    :::|||::::::::::::||||||   |||||||||||    |||
Db     167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
               |||   |||   |||:::|||:::        :::       :::
Db     227 GTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy      35 --------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                                   :::       ||||||     |||
Db     287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy      45 --------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                                      ||||||   |||||||||          |||:::
Db     347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy      60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
                  |||||||||:::|||    |||    :::   |||   ||||||       :::||||||
Db     407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy      80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
               ||||||||||||||::::::::    |||:::||||||||||||||      |||:::
Db     467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy     100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
               |||||||||||||||||:::|||      |||:::   ||||||::::::||||||||||||
Db     527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy     119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
               ||||||   :::   ||||||:::::::::||||||        |||      ||||||
Db     587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy     139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
                  :::||||||||||||||||||||||      |||:::   ||||||::::::      |||
Db     647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy     159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
               ||||||      ::::::::::::   |||   |||   |||||||||   ||||||:::
Db     707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy     179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
               ||||||||||||||||||||||||||||||||:::::||||||||||||||||||||||
Db     767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy     199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
                  |||:::||||||||:::
Db     827 GGACAAAAGAAAGAACTTGAA--------------------------------------- 847

Qy     219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
                                                       ||||||   :::||||||
Db     848 ----------------------------------------AGGTACAACAAAAATTTA 865

Qy     239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
               ::::::||||||      |||||||||:::   |||   :::   :::|||:::||||||
Db     866 GAAGAAGCTAAAGGAATTGGGATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCC 925

Qy     259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
                  :::      :::      :::|||   |||||||||||||||||||||||||:::
```

```
Db     926 GCTTTCTTATTGATCTATGCATCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTC 985

Qy     279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
           |||:::   |||       |||||||||||  ||||||  ||||||||||||||:::
Db     986 CTCTCCAGTGAA------TATACTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTA 1039

Qy     299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
                 ::::::   ||||||  |||   |||      |||  ||||||  ||||||
Db    1040 ATTGGGGCTTTTAGTATTGGACAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGA 1099

Qy     319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
           |||||||||:::   ||||||::::::||||||  ||||||||||||||||:::::::|||
Db    1100 GGAGCAGCTTATGAAATCTTCAAGATAATTGACAATAAACCAAGCATTGACAGCTATTCG 1159

Qy     339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
              :::|||:::||||||::::::|||:::|||   :::||||||||||||||||   |||
Db    1160 AAGAGTGGACATAAACCAGATAATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTC 1219

Qy     359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
           :::||||||||||||      :::|||||||||||||||||||||||||:::::::::|||
Db    1220 AGTTACCCTTCTCGAAAAGAAGTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGT 1279

Qy     379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
           |||:::||||||||||||||||||    :::|||  ||||||||||||   |||||||||
Db    1280 GGGCAGACAGTGGCGCTGGTTGGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTG 1339

Qy     399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
           :::||||||||||||||||||   ||||||   :::   :::|||         ||||||||
Db    1340 ATGCAGAGGCTCTATGACCCCACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGG 1399

Qy     419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
              :::||||||||||||   |||:::   ||||||||||||||||||||||||||||
Db    1400 ACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTG 1459

Qy     439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
           |||  |||||||||:::  ||||||:::|||||||||::::::||||||    :::|||
Db    1460 TTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAG 1519

Qy     459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
           :::|||:::|||   :::|||||||||||||||||||||||||||:::   ||||||||
Db    1520 ATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTTATCATGAAACTACCTAATAAA 1579

Qy     479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
           |||:::||||||||||||||||:::|||||||||:::||||||||||||||||||||||||
Db    1580 TTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATC 1639

Qy     499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
           ||||||||||||||||||||||||||||||||||||:::||||||||||||||||||||
Db    1640 GCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCA 1699

Qy     519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
           |||||||||:::||||||::::::   ||||||  ||||||:::|||
Db    1700 GCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCCCTGGATAAG 1747
```

RESULT 6

```
AAD03504
ID   AAD03504 standard; cDNA; 4279 BP.
XX
AC   AAD03504;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) allelic variant (Genotype A) cDNA.
XX
KW   Dog; P-glycoprotein allelic variant; multidrug transporter; MDR1;
KW   drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag= a
FT                   /product= "Dog P-glycoprotein (PGP) allelic variant
FT                   (Genotype A) protein"
FT   allele          replace (607, C)
FT                   /*tag= b
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00308.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 9; Page 85-90; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) allelic variant
CC   (Genotype A) cDNA. The PGP enzyme functions as an efflux pump
CC   exporting small molecules across the cell membrane. This enzyme
CC   is a member of the ABC transporter family.
XX
SQ   Sequence 4279 BP; 1295 A; 833 C; 1008 G; 1143 T; 0 other;

Alignment Scores:
Pred. No.:              1.81e-150     Length:        4279
Score:                  1557.50       Matches:       303
Percent Similarity:     72.12%        Conservative:  98
Best Local Similarity:  54.50%        Mismatches:    104
Query Match:            56.86%        Indels:        51
DB:                     22            Gaps:          5

US-09-873-409-7 (1-541) x AAD03504 (1-4279)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||::::::|||   :::|||::::::::::::::||||||  |||||||||||||   |||
Db     167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
           |||   |||   |||:::|||:::      :::          :::
Db     227 GTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy      35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                                :::      ||||||      |||
Db     287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy      45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                          ||||||   |||||||||           |||:::
Db     347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy      60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
              |||||||||:::|||   |||   :::   |||   ||||||      :::||||||
```

```
Db    407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy     80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db    467 AAACAATTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy    100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db    527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy    119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db    587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy    139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db    647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy    159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db    707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy    179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db    767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy    199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
Db    827 GGACAAAAGAAAGAACTTGAA--------------------------------------- 847

Qy    219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
Db    848 ------------------------------------------AGGTACAACAAAAATTTA 865

Qy    239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
Db    866 GAAGAAGCTAAAGGAATTGGGATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCC 925

Qy    259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
Db    926 GCTTTCTTATTGATCTATGCATCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTC 985

Qy    279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
Db    986 CTCTCCAGTGAA------TATTCTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTA 1039

Qy    299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
Db   1040 ATTGGGGCTTTTAGTATTGGACAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGA 1099

Qy    319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
Db   1100 GGAGCAGCTTATGAAATCTTCAAGATAATTGACAATAAACCAAGCATTGACAGCTATTCG 1159

Qy    339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
Db   1160 AAGAGTGGACATAAACCAGATAATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTC 1219

Qy    359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
Db   1220 AGTTACCCTTCTCGAAAAGAAGTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGT 1279

Qy    379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
Db   1280 GGGCAGACAGTGGCGCTGGTTGGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTG 1339

Qy    399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
Db   1340 ATGCAGAGGCTCTATGACCCCACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGG 1399

Qy    419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
Db   1400 ACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTG 1459

Qy    439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
Db   1460 TTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAG 1519

Qy    459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
Db   1520 ATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTTATCATGAAACTACCTAATAAA 1579

Qy    479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
Db   1580 TTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATC 1639

Qy    499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
Db   1640 GCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCA 1699

Qy    519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db   1700 GCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCCCTGGATAAG 1747
```

```
RESULT 7
AAD03505
ID   AAD03505 standard; cDNA; 4279 BP.
XX
AC   AAD03505;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.
XX
KW   Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KW   MDR1; drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag= a
FT                   /product= "Dog P-glycoprotein (PGP) allelic variant
FT                   (Genotype B) protein"
FT   allele          replace (91, T)
FT                   /*tag= b
FT   allele          replace (607, C)
FT                   /*tag= c
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00309.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 9; Page 93-99; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
```

```
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) allelic variant
CC   (Genotype B) cDNA. The PGP enzyme functions as an efflux pump
CC   exporting small molecules across the cell membrane. This enzyme
CC   is a member of the ABC transporter family.
XX
SQ   Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;

Alignment Scores:
Pred. No.:                1.81e-150      Length:        4279
Score:                    1557.50        Matches:       303
Percent Similarity:       72.12%         Conservative:  98
Best Local Similarity:    54.50%         Mismatches:    104
Query Match:              56.86%         Indels:        51
DB:                       22             Gaps:          5

US-09-873-409-7 (1-541) x AAD03505 (1-4279)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          |||::::::|||   :::|||::::::::::::::|||||   |||||||||||   |||
Db    167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
          |||   |||   |||:::|||:::          :::      :::
Db    227 GTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy     35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                               :::      ||||||       |||
Db    287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy     45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                         ||||||      |||||||||      |||:::
Db    347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy     60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
             |||||||||:::|||   |||   :::   |||   ||||||      :::||||||
Db    407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy     80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
          ||||||||||||||:::::::::   |||:::||||||||||||||      |||:::
Db    467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy    100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
          ||||||||||||||||||:::|||   |||:::   ||||||::::::||||||||||||
Db    527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy    119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
          ||||||   :::   ||||||:::::::::||||||      |||      ||||||
Db    587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy    139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
             :::||||||||||||||||||||||||   |||:::   ||||||:::::      |||
Db    647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy    159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
          ||||||      ::::::::::::   |||   |||   |||||||||   ||||||:::
Db    707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy    179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
          |||||||||||||||||||||||||||||||::::::||||||||||||||||||
Db    767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy    199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
             |||:::|||||||||:::
Db    827 GGACAAAAGAAAGAACTTGAA--------------------------------------- 847

Qy    219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
                                                      ||||||   :::||||||
Db    848 ------------------------------------------AGGTACAACAAAAATTTA 865

Qy    239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
          ::::::||||||      |||||||||:::   |||   :::   :::|||:::||||||
Db    866 GAAGAAGCTAAAGGAATTGGGATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCC 925

Qy    259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
             :::         :::|||   ||||||||||||||||||||||||||||||:::
Db    926 GCTTTCTTATTGATCTATGCATCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTC 985

Qy    279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
          |||:::   |||      |||:::||||||   ||||||   ||||||||||||:::
Db    986 CTCTCCAGTGAA------TATTCTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTA 1039

Qy    299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
                ::::::   ||||||   |||   |||      |||   ||||||   ||||||
Db   1040 ATTGGGGCTTTTAGTATTGGACAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGA 1099

Qy    319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
          |||||||||:::   ||||||::::::||||||   |||||||||||||||::::::|||
Db   1100 GGAGCAGCTTATGAAATCTTCAAGATAATTGACAATAAACCAAGCATTGACAGCTATTCG 1159

Qy    339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
             :::|||:::||||||::::::|||:::|||   :::|||||||||||||||   |||
Db   1160 AAGAGTGGACATAAACCAGATAATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTC 1219

Qy    359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
             :::|||||||||||||||||||||||||||||||||||||||:::::::::|||
Db   1220 AGTTACCCTTCTCGAAAAGAAGTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGT 1279

Qy    379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
          |||:::|||||||||||||||||||      :::|||   ||||||||||||   |||||||||
Db   1280 GGGCAGACAGTGGCGCTGGTTGGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTG 1339

Qy    399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
          :::|||||||||||||||||||   ||||||   :::   :::|||      |||||||||
Db   1340 ATGCAGAGGCTCTATGACCCCACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGG 1399

Qy    419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
             :::||||||||||||   |||:::      ||||||||||||||||||||||||||||
Db   1400 ACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTG 1459

Qy    439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
          |||   |||   ||||||:::   ||||||:::||||||::::::||||||   :::|||
Db   1460 TTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAG 1519

Qy    459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
          :::|||:::|||   :::||||||||||||||||||||||||||||:::   |||||||||
Db   1520 ATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTTATCATGAAACTACCTAATAAA 1579

Qy    479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
          |||:::|||||||||||||||:::|||||||||:::||||||||||||||||||||||||
Db   1580 TTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATC 1639

Qy    499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
          |||||||||||||||||||||||||||||||||||||||:::||||||||||||||||||
Db   1640 GCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCA 1699

Qy    519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
          |||||||||:::||||||::::::      ||||||   ||||||:::|||
Db   1700 GCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCCCTGGATAAG 1747

RESULT 8
AAF86128
ID   AAF86128 standard; cDNA; 4195 BP.
XX
AC   AAF86128;
XX
DT   25-JUN-2001  (first entry)
XX
DE   Cynomologous monkey P-glycoprotein cDNA variant 1.
XX
```

```
KW   Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   efflux pump; ss.
XX
OS   Macaca fascicularis.
XX
FH   Key             Location/Qualifiers
FT   CDS             100..3951
FT                   /*tag= a
FT                   /product= "PGP"
FT                   /note= "P-glycoprotein"
FT   misc_feature    376..384
FT                   /*tag= b
FT                   /note= "Insertion of 9 nucleotides relative to PGP
FT                          allelic variant AAF86127"
XX
PN   WO200123565-A1.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26592.
XX
PR   28-SEP-1999;   99US-0156921.
PR   12-OCT-1999;   99US-0158818.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-Crespi DT,  Crespi CL;
XX
DR   WPI; 2001-316136/33.
DR   P-PSDB; AAB81065.
XX
PT   Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein
PT   (PGP) and homologous PGP polypeptides are useful for predicting
PT   bioavailability of compound and increasing PGP transporter activity in
PT   cell  -
XX
PS   Example 1; Page 59-65; 84pp; English.
XX
CC   This invention relates to a polynucleotide sequence encoding a
CC   cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the
CC   PGP protein. PGP, also known as multidrug transporter, MDR1 is a member
CC   of the ABC transporter superfamily. The enzyme serves as an efflux pump
CC   exporting small molecules across the cell membrane. The invention
CC   includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence
CC   and protein, and also that of an allelic variant. The PGP polynucleotide
CC   sequence is useful for increasing PGP transporter activity in a cell.
CC   Antisense sequences of the cDNA are useful for inhibiting PGP transport
CC   activity in a mammalian cell. They may also be used for increasing the
CC   bioavailability of a drug. The present sequence represents cDNA encoding
CC   cynomologous monkey P-glycoprotein. This sequence contains a 9 nucleotide
CC   insert compared to the PGP allelic variant given in AAF86127.
XX
SQ   Sequence 4195 BP; 1230 A; 801 C; 1039 G; 1125 T; 0 other;

Alignment Scores:
Pred. No.:              1.99e-150    Length:         4195
Score:                  1557.00      Matches:        306
Percent Similarity:     72.17%       Conservative:   96
Best Local Similarity:  54.94%       Mismatches:     103
Query Match:            56.85%       Indels:         52
DB:                     22           Gaps:           5

US-09-873-409-7 (1-541) x AAF86128 (1-4195)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    250 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGAGCTGGACTTCCTCTCATGATGCTG 309

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSer---------------------GlyCys 33
Db    310 GTGTTTGGAGACATGACGGATACCTTTGCAAATGCAGGAAATTTAGGAGATTTAGGAGCT 369

Qy     34 LeuValGlnThrAsnThrTyrSer------------------------------------ 41
Db    370 CTGTTGTTTAACAACACTAATAGCAGTAATATCACTGATACAGTGCCCGTCATGAATCTG 429

Qy     42 ---------PhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIle 58
Db    430 GAGGAAGATATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTT 489

Qy     59 PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIle 78
Db    490 GCTGCTTACATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATT 549

Qy     79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp 98
Db    550 AGAAAACAGTTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGAT 609

Qy     99 IleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly 117
Db    610 GTTGGGGAGCTTAACACCCGGCTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGT 669

Qy    118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
Db    670 GACAAAATTGGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGA 729

Qy    138 LeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
Db    730 TTTACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTG 789

Qy    158 SerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr 177
Db    790 TCAGCTGCAGTCTGGGCAAAGATACTGTCTTCATTTACTGATAAAGAACTCTTAGCTTAT 849

Qy    178 SerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197
Db    850 GCAAAAGCTGGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTT 909

Qy    198 ArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrp 217
Db    910 GGAGGACAAAAGAAAGAACTCGAA------------------------------------ 933

Qy    218 PheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsn 237
Db    934 ---------------------------------------------AGGTACAACAAAAAT 948

Qy    238 LeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGly 257
Db    949 TTAGAAGAAGCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGT 1008

Qy    258 AlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeu 277
Db   1009 GCTGCTTTCCTGCTTATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTG 1068

Qy    278 IleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerVal 297
Db   1069 GTCCTCTCAAAGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTA 1122

Qy    298 IleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAla 317
Db   1123 TTAATTGGGGCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCA 1182

Qy    318 ArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPhe 337
Db   1183 AGAGGAGCAGCTTTTGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTAT 1242

Qy    338 SerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSer 357
Db   1243 TCGAAGAGTGGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCAC 1302

Qy    358 PheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLys 377
Db   1303 TTCAGTTACCCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAG 1362

Qy    378 SerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGln 397
```

```
Db    1363 AGTGGGCAGACGGTGGCCCTGGTTGGAAACAGCGGCTGTGGGAAGAGCACAACGGTCCAG 1422

Qy     398 LeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIle 417
           |||:::|||||||||||||||||   :::|||   :::   ||||||      ||||||
Db    1423 CTGATGCAGAGGCTTTATGACCCCACAGAGGGCATGGTCAGTGTTGATGGACAGGATATT 1482

Qy     418 ArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProVal 437
           |||   :::||||||||      |||:::  |||||||||||||||||||||||||||
Db    1483 AGGACCATAAACGTAAGGTTTCTACGGGAAATCATCGGTGTGGTGAGTCAGGAACCTGTA 1542

Qy     438 LeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGlu 457
           ||||||  |||||||||:::  ||||||:::||||||||:::|||||||||     :::
Db    1543 TTGTTTGCCACCACGATAGCTGAAAACATTCGCTATGGTCGTGAAGATGTCACCATGGAT 1602

Qy     458 GluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsn 477
           |||:::|||:::|||   :::||||||||||||||||||||||||||||:::    |||
Db    1603 GAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCAG 1662

Qy     478 LysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArg 497
           ||||||:::||||||||||||||||:::|||||||||:::||||||||||||||||||||
Db    1663 AAATTTGACACCCTGGTTGGAGAGAGAGGGGCCCAGCTGAGTGGTGGGCAGAAGCAGAGG 1722

Qy     498 IleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThr 517
           ||||||||||||||||||||||||||||||||||||||||||:::|||||||||||||||
Db    1723 ATCGCCATTGCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGACGAGGCCACG 1782

Qy     518 SerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
           ||||||||||||:::||||||:::::    ||||||   ||||||:::|||
Db    1783 TCAGCCTTGGACACAGAAAGTGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 1833

RESULT 9
AAZ49332
ID   AAZ49332 standard; cDNA; 3860 BP.
XX
AC   AAZ49332;
XX
DT   14-MAR-2000  (first entry)
XX
DE   Human wild-type multidrug resistance-1 (MDR-1) cDNA.
XX
KW   Multidrug resistance; MDR-1; P-glycoprotein;
KW   transmembrane efflux pump; haematopoietic stem cell; transduction;
KW   bone marrow transplantation; chemotherapy; radiation therapy; cancer;
KW   gene therapy; gene replacement; genetic defect; thalassaemia;
KW   Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
KW   cytokine; wild-type; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3843
FT                   /*tag=  a
FT                   /product= "Human wild-type MDR-1 protein"
FT   mutation        replace (553..555, GTT)
FT                   /*tag=  b
FT                   /note= "cDNA sequence of G185V human mutant MDR-1 given
FT                          in AAZ49333"
XX
PN   WO9961589-A2.
XX
PD   02-DEC-1999.
XX
PF   27-MAY-1999;   99WO-US11825.
XX
PR   28-MAY-1998;   98US-0086988.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI   Sorrentino B,  Bunting K;
XX
DR   WPI; 2000-072615/06.
DR   P-PSDB; AAY58186.
XX
PT   Ex vivo expansion of hematopoietic stem cells transduced with a
PT   sequence encoding human multidrug resistance-1, used for bone marrow
PT   transplantation -
XX
PS   Claim 10; Page 68-70; 113pp; English.
XX
CC   This sequence represents cDNA encoding human wild-type
CC   multidrug resistance protein MDR-1. MDR-1 is a transmembrane
CC   efflux pump, responsible for the export of drugs from cells,
CC   particularly cancer cells. Wild-type MDR-1 shows increased
CC   resistance to etoposide and decreased resistance to vinca
CC   alkaloids compared with a mutant form (AAY58187) where the Gly at
CC   position 185 is replaced by Val. The invention relates to transducing
CC   haematopoietic stem cells with nucleic acid encoding an MDR protein
CC   and culturing the modified cells. The modified haematopoietic stem
CC   cells are useful in bone marrow transplantation (to reconstitute
CC   haematopoietic systems in patients who have undergone chemotherapy or
CC   radiation therapy) and in ex vivo gene therapy of genetic defects in
CC   cells derived from haematopoietic stem cells, e.g., thalassaemia,
CC   Gauchier's disease, sickle cell anaemia or leukaemia. The modified
CC   cells can also be used to identify factors involved in regulating
CC   proliferation and differentiation in haematopoietic stem cells.
CC   Haematopoietic stem cells that express MDR-1 will be protected against
CC   chemotherapeutic agents, so can be engrafted while the patient is
CC   undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells
CC   provides sufficient cells to permit standard biochemical analysis.
CC   Overexpression of MDR-1 allows cytokine-driven expansion of
CC   haematopoietic stem cells by at least 10-fold compared with a maximum
CC   of 4-fold in known procedures.
XX
SQ   Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;

Alignment Scores:
Pred. No.:               2.52e-150     Length:        3860
Score:                   1555.50       Matches:       306
Percent Similarity:      73.29%        Conservative:  100
Best Local Similarity:   55.23%        Mismatches:    99
Query Match:             56.79%        Indels:        49
DB:                      21            Gaps:          6

US-09-873-409-7 (1-541) x AAZ49332 (1-3860)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||::::::|||  ||||||::::::::::::||||||   ||||||||||||   |||
Db     151 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210

Qy      21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
           |||   |||||||||:::|||                           :::|||:::|||
Db     211 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 270

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                    ::::::      |||      ||||||
Db     271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                 |||         ||||||   |||||||||      |||:::      |||
Db     331 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 390

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
           ||||||:::|||   |||   :::   |||   ||||||       :::||||||||||||
Db     391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
           ||||||||||:::::::::   |||:::|||||||||||||||      |||:::||||||
Db     451 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 510

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
           ||||||||||||:::|||   |||:::   ||||||::::::|||||||||||||||||||
Db     511 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGGTATTGGTGACAAAATT 570

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
              :::   ||||||:::|||:::||||||      |||         ||||||      :::
Db     571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 630
```

```
Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
           ||||||||||||||||||||      |||:::    ||||||::::::      ||||||||
Db     631 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
                 :::::::::::    |||   |||   ||||||||    ||||||:::||||||
Db     691 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
           ||||||||||||||||||||||||::::::||||||||||||||||||||      |||
Db     751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810

Qy     201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
           :::|||||||||:::
Db     811 AAGAAAGAACTTGAA--------------------------------------------- 825

Qy     221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                              ||||||   :::||||||::::::
Db     826 -----------------------------------AGGTACAACAAAAATTTAGAAGAA 849

Qy     241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
           ||||||   |||||||||:::   |||   :::   :::|||:::||||||     :::
Db     850 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 909

Qy     261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
                 :::      :::|||   ||||||||||||||||||||:::|||:::|||:::
Db     910 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 969

Qy     281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
           ||||||        |||:::||||||    ||||||   ||||||||||||||:::
Db     970 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1023

Qy     301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
           ::::::   :::|||   |||   |||      |||   ||||||   ||||||||||||
Db    1024 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1083

Qy     321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
           |||:::   ||||||::::::||||||   |||||||||||||||::::::|||   :::
Db    1084 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1143

Qy     341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
           |||:::||||||::::::|||:::|||   :::||||||:::||||||   |||:::|||
Db    1144 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1203

Qy     361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
           |||||||||      :::||||||||||||||||||||||||:::::::::||||||:::
Db    1204 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1263

Qy     381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
           ||||||||||||||||||      :::|||   ||||||      ||||||||||:::|||
Db    1264 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1323

Qy     401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
           |||||||||||||||   :::|||   :::   ||||||      |||||||||   :::
Db    1324 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1383

Qy     421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
           |||||||||      |||:::   ||||||||||||||||||||||||||||||||
Db    1384 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1443

Qy     441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
           |||||||||:::   ||||||:::|||||||||::::::||||||   :::|||:::|||
Db    1444 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1503

Qy     461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
           :::|||   :::||||||||||||||||||||||||:::   |||:::|||||||:::
Db    1504 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1563

Qy     481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
           ||||||||||||||:::||||||||:::||||||||||||||||||||||||||||||
Db    1564 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 1623

Qy     501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
           ||||||||||||||||||||||||||||||:::||||||||||||||||||||||||
Db    1624 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 1683

Qy     521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
           |||:::||||||::::::   ||||||   ||||||:::|||
Db    1684 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 1725
```

```
RESULT 10
ABA94365
ID   ABA94365 standard; DNA; 3860 BP.
XX
AC   ABA94365;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human BCRP DNA related seq Id No. 1.
XX
KW   Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;
KW   hepatotropic; neurodegenerative; cytostatic; antianemic; muscular; BCRP;
KW   cardiant; gene therapy; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3843
FT                   /*tag=  a
XX
PN   WO200192877-A2.
XX
PD   06-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US17459.
XX
PR   31-MAY-2000; 2000US-0584586.
PR   29-MAY-2001; 2001US-0866866.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI   Sorrentino B,  Schuetz J;
XX
DR   WPI; 2002-114368/15.
DR   P-PSDB; ABB07266.
XX
PT   Identifying a stem cell, for treating e.g., muscular dystrophy,
PT   myocardial infarction, Parkinson's disease, or neurodegenerative
PT   disorders, comprises detecting the expression of an ATP transport
PT   protein (BCRP) by a cell  -
XX
PS   Disclosure; Page 53-55; 87pp; English.
XX
CC   The invention provides a method of identifying and/or isolating a stem
CC   cell that involves detecting the expression of an ATP transport protein
CC   containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
CC   comprising stem cells. The isolated stem cells may be used in the
CC   treatment of diseases such as muscular dystrophy, degenerative liver
CC   disorder, myocardial infarction, Parkinson's disease, degenerative
CC   disorders of the brain, and for tissue regeneration or replacement.
CC   Haematopoietic cells can be used in bone marrow transplants (e.g., for
CC   treatment of leukemia) and for ex vivo gene therapy for treating blood
CC   diseases such as sickle cell anemia and thalassemia. The stem cells can
CC   also be used as cell targets in gene therapy protocols. The present
CC   sequence represents a sequence related to the BCRP for which no relevant
CC   information has been provided in the specification.
XX
SQ   Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;

Alignment Scores:
Pred. No.:              2.52e-150     Length:         3860
Score:                  1555.50       Matches:        306
Percent Similarity:     73.29%        Conservative:   100
Best Local Similarity:  55.23%        Mismatches:     99
Query Match:            56.79%        Indels:         49
```

```
DB:                    24          Gaps:        6

US-09-873-409-7 (1-541) x ABA94365 (1-3860)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    151 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210

Qy     21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db    211 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 270

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db    271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    331 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 390

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    451 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 510

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    511 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGGTATTGGTGACAAAATT 570

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 630

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    631 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    691 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db    811 AAGAAAGAACTTGAA--------------------------------------------- 825

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db    826 ------------------------------------AGGTACAACAAAAATTTAGAAGAA 849

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db    850 GCTAAAAGAATTGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 909

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db    910 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 969

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db    970 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1023

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db   1024 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1083

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   1084 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1143

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   1144 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1203

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   1204 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1263

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   1264 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1323

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   1324 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1383

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   1384 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1443

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   1444 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1503

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   1504 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1563

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   1564 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 1623

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   1624 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 1683

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db   1684 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 1725

RESULT 11
AAZ90198
ID   AAZ90198 standard; DNA; 4233 BP.
XX
AC   AAZ90198;
XX
DT   19-MAY-2000  (first entry)
XX
DE   Rat mdr1b2 (multispecific drug transporter) nucleotide sequence.
XX
KW   Rat; mdr1b2; multispecific drug transporter; cancer; inflammation;
KW   cardiovascular disease; central nervous system disorder;
KW   autoimmune disease; kidney disease; drug formulation; ss.
XX
OS   Rattus sp.
XX
PN   US6025160-A.
XX
PD   15-FEB-2000.
XX
PF   22-JUL-1998;   98US-0120513.
XX
PR   22-JUL-1998;   98US-0120513.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
XX
PI   Brun KA,  Ellens H,  Yue L,  Chenery RJ,  Feild JA;
XX
DR   WPI; 2000-181810/16.
DR   P-PSDB; AAY78879.
```

```
XX
PT   Isolated polynucleotide encoding a rat mdr1b2 drug transporter
PT   polypeptide, useful for treatment of e.g. cancer, autoimmune disease,
PT   central nervous system disorders  -
XX
PS   Claim 7; Column 13-17; 17pp; English.
XX
CC   This sequence represents a polynucleotide encoding a 1275 amino acid
CC   rat mdr1b2 multispecific drug transporter polypeptide. The mdr1b2
CC   nucleotide sequence has cytostatic, antiinflammatory, cardiant,
CC   neuroprotective, immunosuppressive and nephrotropic activity.
CC   Understanding the functioning of the mdr1b2 polynucleotide and protein in
CC   transgenic animal models is useful for treating and preventing diseases
CC   such as cancer, inflammation, cardiovascular disease, central nervous
CC   system disorders, autoimmune disease, and kidney disease. The use of the
CC   protein in cell based, membrane based, or binding assays may enhance drug
CC   formulation, selection of formulation excipients and compound design.
XX
SQ   Sequence 4233 BP; 1210 A; 920 C; 1048 G; 1055 T; 0 other;

Alignment Scores:
Pred. No.:              5.88e-150     Length:         4233
Score:                  1552.50       Matches:        308
Percent Similarity:     72.20%        Conservative:   92
Best Local Similarity:  55.60%        Mismatches:     105
Query Match:            56.68%        Indels:         49
DB:                     21            Gaps:           5

US-09-873-409-7 (1-541) x AAZ90198 (1-4233)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    170 ATGGCTCTGGGAACTCTCGCTGCTATCATCCACGGAACCCTGCTTCCCCTCCTGATGCTG 229

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
Db    230 GTGTTCGGATACATGACAGATAGTTTACCCAAGCAGAGACCCGCATTCTGCCGAGCGTT 289

Qy     31 ------SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu--------- 45
Db    290 ACTAATCAAAGTGAAATCAACAGTACACAGACCGTCAGCGACAGCAGTCTGGAGGAGGAC 349

Qy     46 ------------ThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    350 ATGGCCATGTACGCCTACTATTACACGGGCATTGGTGCCGGTGTGCTCATCGTTGCCTAC 409

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    410 ATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAAG 469

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    470 TTTTTCCATGCCATCATGAATCAGGAGATAGGCTGGTTTGACGTGAATGACGCTGGGGAG 529

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    530 CTCAACACCCGGCTCACAGATGACGTCTCCAAAATTAATGACGGAATTGGTGACAAACTT 589

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    590 GGAATGTTCTTTCAGTCCATAACGACATTTTCAGCCGGTTTTATAATAGGATTTATAAGT 649

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    650 GGTTGGAAGCTAACCCTTGTAATTTTGGCCGTCAGCCCTCTTATTGGGTTGTCATCTGCC 709

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    710 ATGTGGGCAAAGGTACTGACTTCATTTACTAATAAGGAACTCCAGGCTTATGCGAAAGCT 769

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    770 GGAGCAGTTGCCGAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCGTTTGGAGGACAA 829

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db    830 AAGAAGGAACTTGAA--------------------------------------------- 844

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db    845 ------------------------------------AGGTACAATAAAAATTTAGAAGAA 868

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db    869 GCTAAAAGAGTTGGCATAAAGAAAGCCATCACGGCCAACATTTCCATAGGTATTGCCTAC 928

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db    929 CTGTTGGTCTATGCGTCTTATGCACTGGCATTCTGGTATGGGACCTCCTTGGTCCTCTCA 988

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db    989 AATGAA------TATTCTATTGGACAAGTGCTTACCGTCTTCTTCTCTATTTTATTGGGG 1042

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db   1043 ACTTTCAGTATTGGACATTTAGCCCCAAACATAGAAGCCTTTGCAAATGCAAGAGGGGCA 1102

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   1103 GCCTATGAAATCTTCAAGATAATTGATAATGAGCCAAGCATTGACAGCTTCTCAACCAAG 1162

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   1163 GGACACAAACCAGACAGTATAATGGGAAATTTGGAATTTAAAAATGTTTACTTCAACTAC 1222

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   1223 CCATCACGAAGTGAAGTTAAGATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAG 1282

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   1283 ACGGTAGCCCTGGTTGGCAACAGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAG 1342

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   1343 AGGCTCTACGACCCCATAGAGGGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATC 1402

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   1403 AATGTGAGGTATCTGCGGGAAATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCC 1462

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   1463 ACCACGATTGCCGAAAACATTCGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAG 1522

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   1523 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTCATCATGAAACTGCCCCACAAATTTAAC 1582

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   1583 ACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATT 1642

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   1643 GCCCGGGCCCTGGTCCGCAACCCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTG 1702

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db   1703 GACACAGAAAGCGAAGCCGTGGTTCAGGCCGCTCTGGATAAG 1744

RESULT 12
AAF27498
ID   AAF27498 standard; cDNA; 4233 BP.
XX
```

```
AC   AAF27498;
XX
DT   25-APR-2001  (first entry)
XX
DE   Rat mdr1b2 multidrug resistance cDNA, SEQ ID NO:1.
XX
KW   Rat; multidrug resistance; mdr1b2; multi-drug transporter family;
KW   drug screening; pharmacokinetic analysis; oral absorption;
KW   formulation design; bioavailability; transgenic animal; knockout animal;
KW   inflammation; cardiovascular disease; central nervous system disorder;
KW   cancer; autoimmune disorder; kidney disease; ss.
XX
OS   Rattus sp.
XX
PN   US6169166-B1.
XX
PD   02-JAN-2001.
XX
PF   29-NOV-1999;   99US-0450105.
XX
PR   22-JUL-1998;   98US-0120513.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI   Chenery RJ,  Ellens H,  Brun KA,  Yue L,  Feild JA;
XX
DR   WPI; 2001-158280/16.
DR   P-PSDB; AAB60409.
XX
PT   New rat multidrug resistance protein (mdr1b2), useful for screening or
PT   identifying compounds that are (ant)agonists/inhibitors of the
PT   ratmdr1b2, as well as compounds with optimal development
PT   characteristics  -
XX
PS   Disclosure; Column 13-14; 13pp; English.
XX
CC   The invention relates to a novel rat multidrug resistance (mdr)
CC   protein, mdr1b2 (AAB60409). Multi-specific drug transporter family
CC   proteins are present in cell which have a barrier function, such as
CC   intestinal epithelial cells, brain microvessel endothelial cells,
CC   kidney epithelial cells, and liver hepatocytes, and are also expressed
CC   by certain cancer cells. The rat mdr1b2 protein is useful for
CC   screening or identifying compounds that are agonists or antagonists
CC   of mdr1b2 activity. It may also be used to establish assays to
CC   predict oral absorption and pharmacokinetics of drugs in humans, and
CC   thus enhance the design of formulations through the identification of
CC   compounds with optimal development characteristics (i.e., high oral
CC   bioavailability, UID (once a day) dosing, reduced drug interactions,
CC   reduced variability, and reduced food effects), specifically to
CC   avoid interactions with human mdr-1. Transgenic and knockout animals
CC   created using DNA encoding the rat mdr1b2 may be used to gain an
CC   insight into treating and preventing human diseases such as cancer,
CC   inflammation, cardiovascular disease, central nervous system disorders,
CC   autoimmune disorders and kidney disease. The present sequence represents
CC   cDNA encoding rat mdr1b2.
XX
SQ   Sequence 4233 BP; 1210 A; 920 C; 1048 G; 1055 T; 0 other;

Alignment Scores:
Pred. No.:              5.88e-150     Length:          4233
Score:                  1552.50       Matches:         308
Percent Similarity:     72.20%        Conservative:    92
Best Local Similarity:  55.60%        Mismatches:      105
Query Match:            56.68%        Indels:          49
DB:                     22            Gaps:            5

US-09-873-409-7 (1-541) x AAF27498 (1-4233)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
             |||   ||||||   ||||||::::::::::::|||       |||||||||:::   |||
Db     170 ATGGCTCTGGGAACTCTCGCTGCTATCATCCACGGAACCCTGCTTCCCCTCCTGATGCTG 229

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
           |||   |||   ||||:::||||:::
Db     230 GTGTTCGGATACATGACAGATAGTTTTACCCAAGCAGAGACCCGCATTCTGCCGAGCGTT 289

Qy      31 ------SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu--------- 45
                      :::         |||   |||   |||         |||
Db     290 ACTAATCAAAGTGAAATCAACAGTACACAGACCGTCAGCGACAGCAGTCTGGAGGAGGAC 349

Qy      46 -----------ThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                               |||||   ||||||||         |||||      |||
Db     350 ATGGCCATGTACGCCTACTATTACACGGGCATTGGTGCCGGTGTGCTCATCGTTGCCTAC 409

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
           ||||||:::||||||||     :::    |||   ||||||        :::||||||::::::
Db     410 ATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAAG 469

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
           ||||||||:::::::::   |||:::|||||||||||||          |||   ||||||
Db     470 TTTTTCCATGCCATCATGAATCAGGAGATAGGCTGGTTTGACGTGAATGACGCTGGGGAG 529

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
           |||||||||||:::|||     |||:::      ||||||:::||||||||||||||||||:::
Db     530 CTCAACACCCGGCTCACAGATGACGTCTCCAAAATTAATGACGGAATTGGTGACAAACTT 589

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
              :::    ||||||:::::::::||||||||   |||        :::|||    :::
Db     590 GGAATGTTCTTTCAGTCCATAACGACATTTTCAGCCGGTTTTATAATAGGATTTATAAGT 649

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
           ||||||||||||||||||||     |||:::   ||||||||||||      |||:::|||
Db     650 GGTTGGAAGCTAACCCTTGTAATTTGGCCGTCAGCCCTCTTATTGGGTTGTCATCTGCC 709

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
                 ::::::::::::    |||   |||:::||||||||   ||||||:::||||||
Db     710 ATGTGGGCAAAGGTACTGACTTCATTTACTAATAAGGAACTCCAGGCTTATGCGAAAGCT 769

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
           ||||||||||||||||||||||||:::::|||||||||||||||||||||||||||
Db     770 GGAGCAGTTGCCGAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCGTTTGGAGGACAA 829

Qy     201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
           :::|||||||||:::
Db     830 AAGAAGGAACTTGAA--------------------------------------------- 844

Qy     221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                              ||||||   :::||||||::::::
Db     845 -----------------------------------AGGTACAATAAAAATTTAGAAGAA 868

Qy     241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
           ||||||      |||||||||:::   |||    :::   :::|||:::|||       |||
Db     869 GCTAAAAGAGTTGGCATAAAGAAAGCCATCACGGCCAACATTTCCATAGGTATTGCCTAC 928

Qy     261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
              :::        :::|||   ||||||||||||||||||||||||||||:::|||:::
Db     929 CTGTTGGTCTATGCGTCTTATGCACTGGCATTCTGGTATGGGACCTCCTTGGTCCTCTCA 988

Qy     281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPheSerValIleHisSer 300
              |||      |||:::|||||||   ||||||   |||||||||||||::::::
Db     989 AATGAA------TATTCTATTGGACAAGTGCTTACCGTCTTCTTCTCTATTTATTGGGG 1042

Qy     301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
           ::::::   ||||||         |||:::   |||   ||||||   ||||||||||||
Db    1043 ACTTTCAGTATTGGACATTTAGCCCCAAACATAGAAGCCTTTGCAAATGCAAGAGGGGCA 1102

Qy     321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
           |||:::   ||||||::::::||||||      :::|||||||||||:::|||||||||
Db    1103 GCCTATGAAATCTTCAAGATAATTGATAATGAGCCAAGCATTGACAGCTTCTCAACCAAG 1162

Qy     341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
           |||:::||||||:::||||||   |||      :::||||||||||||||    |||||||||
Db    1163 GGACACAAACCAGACAGTATAATGGGAAATTTGGAATTTAAAAATGTTTACTTCAACTAC 1222

Qy     361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
```

```
               ||||||||       :::||||||||||||||||||||||||::::::|||||||||:::
Db    1223 CCATCACGAAGTGAAGTTAAGATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAG 1282

Qy     381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
           ||||||||||||||||||    :::|||  |||||||||||    ||||||||||||||
Db    1283 ACGGTAGCCCTGGTTGGCAACAGTGGCTGTGGAAAAGCACAACTGTCCAGCTGCTGCAG 1342

Qy     401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
           ||||||||||||||   :::|||   :::    :::|||        ||||||||   :::
Db    1343 AGGCTCTACGACCCCATAGAGGGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATC 1402

Qy     421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
           |||||||||:::   |||:::  ||||||||||||||||||||||||||||||
Db    1403 AATGTGAGGTATCTGCGGGAAATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCC 1462

Qy     441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
           |||||||||:::   ||||||:::||||||||||::::::||||||    :::|||:::|||
Db    1463 ACCACGATTGCCGAAAACATTCGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAG 1522

Qy     461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
           :::|||   :::|||||||||||||||||||||||||||||||:::   |||:::|||||||||
Db    1523 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTCATCATGAAACTGCCCCACAAATTTAAC 1582

Qy     481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
           |||||||||||||||:::|||||||||||||||:::|||||||||||||||||||||||||
Db    1583 ACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATT 1642

Qy     501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
           ||||||||||||||||||||||||||||||||:::||||||||||||||||||||||||||
Db    1643 GCCCGGGCCCTGGTCCGCAACCCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTG 1702

Qy     521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
           |||:::||||||:::::   ||||||||||||:::|||
Db    1703 GACACAGAAAGCGAAGCCGTGGTTCAGGCCGCTCTGGATAAG 1744
```

RESULT 13
AAD03488

```
ID   AAD03488 standard; cDNA; 4279 BP.
XX
AC   AAD03488;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) cDNA #1.
XX
KW   Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag=  a
FT                   /product= "Dog P-glycoprotein (PGP) #1"
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00303.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 3; Page 58-63; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) cDNA. This
CC   sequence is also referred as Genotype C cDNA. The
CC   PGP enzyme functions as an efflux pump exporting small molecules
CC   across the cell membrane. This enzyme is a member of the ABC
CC   transporter family.
XX
SQ   Sequence 4279 BP; 1294 A; 834 C; 1008 G; 1143 T; 0 other;

Alignment Scores:
Pred. No.:              5.97e-150       Length:         4279
Score:                  1552.50         Matches:        302
Percent Similarity:     71.94%          Conservative:   98
Best Local Similarity:  54.32%          Mismatches:     105
Query Match:            56.68%          Indels:         51
DB:                     22              Gaps:           5

US-09-873-409-7 (1-541) x AAD03488 (1-4279)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||::::::|||    :::|||::::::::::::||||||    |||||||||||    |||
Db     167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
           |||   |||   |||:::|||:::        :::       :::
Db     227 GTTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy      35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                                :::      ||||||   |||
Db     287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy      45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                                ||||||   |||||||||            |||:::
Db     347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy      60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
              |||||||||:::|||   |||    :::   |||   ||||||       :::||||||
Db     407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy      80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
           |||||||||||||||:::::::::   |||:::||||||||||||||||       |||:::
Db     467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy     100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
           ||||||||||||||||||:::|||   |||:::   ||||||::::::||||||||||||
Db     527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy     119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
           ||||||   :::   |||   ::::::::||||||   |||         ||||||
Db     587 AAAATTGGAATGTTCTTTCACTCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy     139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
              :::|||||||||||||||||||||   |||:::   ||||||::::::       |||
Db     647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy     159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
           ||||||      :::::::::::   |||   |||   |||||||||   ||||||:::
Db     707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766
```

```
Qy     179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
           ||||||||||||||||||||||||||||||||::::::|||||||||||||||||||||||
Db     767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy     199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
              |||:::|||||||||:::
Db     827 GGACAAAAGAAAGAACTTGAA--------------------------------------- 847

Qy     219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
                                                 ||||||   :::||||||
Db     848 -----------------------------------------AGGTACAACAAAAATTTA 865

Qy     239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
           ::::::||||||       |||||||||:::   |||    :::    :::|||:::||||||
Db     866 GAAGAAGCTAAAGGAATTGGGATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCC 925

Qy     259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
              :::       :::       :::|||   ||||||||||||||||||||||||||:::
Db     926 GCTTTCTTATTGATCTATGCATCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTC 985

Qy     279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
           |||:::   |||       |||:::||||||   ||||||   ||||||||||||||:::
Db     986 CTCTCCAGTGAA------TATTCTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTA 1039

Qy     299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
                 ::::::   ||||||   |||   |||      |||   ||||||   ||||||
Db    1040 ATTGGGGCTTTTAGTATTGGACAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGA 1099

Qy     319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
           |||||||||:::   ||||||::::::||||||   ||||||||||||||||::::::|||
Db    1100 GGAGCAGCTTATGAAATCTTCAAGATAATTGACAATAAACCAAGCATTGACAGCTATTCG 1159

Qy     339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
              :::|||:::||||||||||||||||||::::||||   :::||||||||||||||   |||
Db    1160 AAGAGTGGACATAAACCAGATAATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTC 1219

Qy     359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
           :::||||||||||||         :::||||||||||||||||||||||:::::::::|||
Db    1220 AGTTACCCTTCTCGAAAAGAAGTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGT 1279

Qy     379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
           |||:::||||||||||||||||||   :::|||   ||||||||||||   |||||||||
Db    1280 GGGCAGACAGTGGCGCTGGTTGGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTG 1339

Qy     399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
           :::|||||||||||||||||||   ||||||   :::   :::|||      |||||||||
Db    1340 ATGCAGAGGCTCTATGACCCCACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGG 1399

Qy     419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
              :::||||||||||||   |||:::   |||||||||||||||||||||||||||||
Db    1400 ACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTG 1459

Qy     439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
           |||   |||||||||:::   ||||||:::||||||||::::::||||||   :::|||
Db    1460 TTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAG 1519

Qy     459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
           :::|||:::|||   :::|||||||||||||||||||||||||:::   |||||||||
Db    1520 ATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTTATCATGAAACTACCTAATAAA 1579

Qy     479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
           |||:::||||||||||||:::|||||||||:::|||||||||||||||||||||||||||
Db    1580 TTTGACACTCTGGTTGGAGAGAGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATC 1639

Qy     499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
           ||||||||||||||||||||||||||||||||||||:::|||||||||||||||||||
Db    1640 GCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCA 1699

Qy     519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
           |||||||||:::||||||::::::   ||||||   ||||||:::|||
Db    1700 GCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCCCTGGATAAG 1747
```

```
RESULT 14
AAD03489
ID   AAD03489 standard; cDNA; 4317 BP.
XX
AC   AAD03489;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) cDNA #2.
XX
KW   Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             70..3912
FT                   /*tag=  a
FT                   /product= "Dog P-glycoprotein (PGP) #2"
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00304.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 1; Page 66-72; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) cDNA. The
CC   PGP enzyme functions as an efflux pump exporting small molecules
CC   across the cell membrane. This enzyme is a member of the ABC
CC   transporter family.
XX
SQ   Sequence 4317 BP; 1293 A; 844 C; 1019 G; 1161 T; 0 other;

Alignment Scores:
Pred. No.:              6.04e-150     Length:         4317
Score:                  1552.50       Matches:        302
Percent Similarity:     71.94%        Conservative:   98
Best Local Similarity:  54.32%        Mismatches:     105
Query Match:            56.68%        Indels:         51
DB:                     22            Gaps:           5

US-09-873-409-7 (1-541) x AAD03489 (1-4317)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            |||::::::|||   :::|||::::::::::::||||||   ||||||||||||   |||
```

```
Db     217 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 276

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu----------------- 34
           |||   |||   |||:::|||:::      :::     :::
Db     277 GTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 336

Qy      35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                                :::      ||||||      |||
Db     337 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 396

Qy      45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                                 ||||||     ||||||||          |||:::
Db     397 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 456

Qy      60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
                 ||||||||:::|||    |||   :::    |||   ||||||       :::||||||
Db     457 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 516

Qy      80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
           ||||||||||||||:::::::::     |||:::||||||||||||||       |||:::
Db     517 AAACAATTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 576

Qy     100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
           ||||||||||||||||||:::|||      |||:::   ||||||:::::||||||||||||
Db     577 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 636

Qy     119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
           |||:::    :::       ||||||:::::::::||||||       |||     ||||||
Db     637 AAAGTTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 696

Qy     139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
                 ||||||||||||||||||||||     |||:::     ||||||:::::          |||
Db     697 ACACCTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 756

Qy     159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
           ||||||        ::::::::::::    |||     |||     |||||||||     ||||||:::
Db     757 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 816

Qy     179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
           |||||||||||||||||||||||||||||||:::::::||||||||||||||||||||
Db     817 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 876

Qy     199 AlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPhe 218
              |||:::|||||||||:::
Db     877 GGACAAAAGAAAGAACTTGAA--------------------------------------- 897

Qy     219 TyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeu 238
                                                           ||||||    :::||||||
Db     898 ------------------------------------------AGGTACAACAAAAATTTA 915

Qy     239 LysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAla 258
           ::::::||||||        |||||||||:::     |||     :::     :::|||:::||||||
Db     916 GAAGAAGCTAAAAGAATTGGGATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCC 975

Qy     259 ValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIle 278
                 :::        :::        :::|||   |||||||||||||||||||||||||:::
Db     976 GCTTTCTTATTGATCTATGCATCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTC 1035

Qy     279 LeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIle 298
           |||:::   |||        ||||||||||||   ||||||   |||||||||||||||:::
Db    1036 CTCTCCAGTGAA------TATACTATTGGACAGGTACTCACTGTCTTCTTTTCTGTATTA 1089

Qy     299 HisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArg 318
                 ::::::      ||||||      |||     |||      ||||||     ||||||
Db    1090 ATTGGGGCTTTTAGTATTGGACAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGA 1149

Qy     319 GlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
           |||||||||:::      |||||:::::::|||||      |||||||||||||||:::::::|||
Db    1150 GGAGCAGCTTATGAAATCTTCAAGATAATTGACAATAAACCAAGCATTGACAGCTATTCG 1209

Qy     339 ThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPhe 358
                 :::|||:::||||||::::::|||:::|||     :::||||||||||||||     |||
Db    1210 AAGAGTGGACATAAACCAGATAATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTC 1269

Qy     359 AsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSer 378
           :::||||||||||||       :::|||||||||||||||||||||||:::::::::|||
Db    1270 AGTTACCCTTCTCGAAAAGAAGTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGT 1329

Qy     379 GlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeu 398
           |||:::||||||||||||||||||    :::|||   ||||||||||||   |||||||||
Db    1330 GGGCAGACAGTGGCGCTGGTTGGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTG 1389

Qy     399 LeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArg 418
           :::||||||||||||||||||   ||||||   :::    :::|||       |||||||||
Db    1390 ATGCAGAGGCTCTATGACCCCACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGG 1449

Qy     419 AlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeu 438
              :::||||||||||||   |||:::            ||||||||||||||||||||||||
Db    1450 ACCATAAATGTAAGGCATCTTCGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTG 1509

Qy     439 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 458
           |||   |||||||||:::   ||||||:::||||||||||:::::::||||||   :::|||
Db    1510 TTTGCCACCACGATAGCTGAAAACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAG 1569

Qy     459 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 478
           :::|||:::|||   :::||||||||||||||||||||||||||||:::     |||||||||
Db    1570 ATTGAGAAAGCTGTTAAGGAAGCCAATGCCTATGATTTTATCATGAAACTACCTAATAAA 1629

Qy     479 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 498
           |||:::||||||||||||||||:::||||||::::::||||||||||||||||||||||||
Db    1630 TTTGACACTCTGGTTGGAGAGAGAGGGGCCCGGCTGAGTGGTGGACAGAAACAGAGAATC 1689

Qy     499 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 518
           ||||||||||||||||||||||||||||||||||:::|||||||||||||||||||
Db    1690 GCCATTGCTCGGGCCCTGGTTCGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCA 1749

Qy     519 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
           ||||||||:::||||||::::::    ||||||     |||||:::|||
Db    1750 GCTCTGGACACTGAAAGTGAAGCAGTGGTTCAGGTGGCCCTGGATAAG 1797
```

RESULT 15
AAV32645

```
ID   AAV32645 standard; cDNA; 4669 BP.
XX
AC   AAV32645;
XX
DT   23-SEP-1998  (first entry)
XX
DE   Human P glycoprotein (Pgp) cDNA.
XX
KW   Human P glycoprotein; Pgp; multi-drug resistance; cancer;
KW   UIC2 monoclonal antibody; mAB; cytotoxic; transmembrane efflux pump; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1..424
FT                   /*tag=  a
FT   CDS             425..4267
FT                   /*tag=  b
FT                   /product= "Wild-type Pgp"
FT   3'UTR           4268..4669
FT                   /*tag=  c
XX
PN   WO9821325-A1.
XX
PD   22-MAY-1998.
XX
PF   17-NOV-1997;   97WO-US21214.
XX
PR   15-NOV-1996;   96US-0752447.
XX
PA   (INGE-) INGENEX INC.
PA   (UNII ) UNIV ILLINOIS FOUND.
XX
PI   Mechetner E,  Roninson IB;
```

```
XX
DR   WPI; 1998-297930/26.
DR   P-PSDB; AAW48997.
XX
PT   Immunological reagent specific for P-glyco:protein - useful for
PT   detecting multi-drug resistant cancer, isolating haematopoietic
PT   cells and selective cell killing
XX
PS   Claim 1; Fig 1A; 89pp; English.
XX
CC   The present sequence represents the human P glycoprotein (Pgp)
CC   cDNA which encodes the wild-type Pgp protein.  Pgp is a transmembrane
CC   efflux pump protein involved in multi-drug resistance of cancer cells.
CC   The invention provides methods for developing and using immunological
CC   reagents specific for certain mutant forms of Pgp and wild-type Pgp in
CC   a conformation associated with substrate binding or in the presence
CC   of ATP depleting agents.  An example of the immunological reagent
CC   is the UIC2 monoclonal antibody (mAB).  mAB UIC2 specifically binds
CC   to Pgp in a particular biochemical conformation and is capable of
CC   inhibiting drug efflux from Pgp-expressing cells.  The immunological
CC   reagents are claimed to be useful for detecting Pgp expression in
CC   mammalian cells, including low level expression, particularly in cancer
CC   cells to diagnose multi-drug resistance.  The invention claims that
CC   these immunological reagents are more specific than known reagents for
CC   detecting Pgp and they also eliminate the need for costly and laborious
CC   screening of Pgp inhibitors by cytotoxicity or dye exclusion methods.
XX
SQ   Sequence 4669 BP; 1393 A; 894 C; 1130 G; 1252 T; 0 other;

Alignment Scores:
Pred. No.:              6.77e-150     Length:        4669
Score:                  1552.50       Matches:       305
Percent Similarity:     73.29%        Conservative:  101
Best Local Similarity:  55.05%        Mismatches:    99
Query Match:            56.68%        Indels:        49
DB:                     19            Gaps:          6

US-09-873-409-7 (1-541) x AAV32645 (1-4669)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy     21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db    635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db    695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    875 TTTTTTCATGCTATTATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    935 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAAATT 994

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db   1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db   1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db   1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db   1235 AAGAAAGAACTTGAA--------------------------------------------- 1249

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db   1250 -----------------------------------AGGTACAACAAAAATTTAGAAGAA 1273

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db   1274 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 1333

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db   1334 CTGCTGATCTATGCATCTTATGCTGTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 1393

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db   1394 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1447

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db   1448 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1507

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   1508 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1567

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   1568 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1627

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   1628 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1687

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   1688 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1747

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   1748 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1807

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   1808 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1867

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   1868 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1927

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   1928 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1987

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   1988 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 2047

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   2048 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 2107
```

```
Qy     521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
           |||:::||||||::::::   ||||||   ||||||:::|||
Db    2108 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 2149
```

Search completed: March 30, 2003, 03:21:35
Job time : 587.737 secs

```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 30, 2003, 03:22:44 ; Search time 3721.1 Seconds
                                           (without alignments)
                                           2354.618 Million cell updates/sec

Title:          US-09-873-409-7
Perfect score:  2739
Sequence:       1 MILGILASLVNGACLPLMPL..........SESKSAVQAALEKDTPRYSF 541

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters:      32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115420_19240/app_query.fasta_1.7544
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_30544_@runat_27032003_115420_19240 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      EST:*
                1:  em_estba:*
                2:  em_esthum:*
                3:  em_estin:*
                4:  em_estmu:*
                5:  em_estov:*
                6:  em_estpl:*
                7:  em_estro:*
                8:  em_htc:*
                9:  gb_est1:*
                10:  gb_est2:*
                11:  gb_htc:*
                12:  gb_est3:*
                13:  gb_est4:*
                14:  gb_est5:*
                15:  em_estfun:*
                16:  em_estom:*
                17:  gb_gss:*
                18:  em_gss_hum:*
                19:  em_gss_inv:*
                20:  em_gss_pln:*
                21:  em_gss_vrt:*
                22:  em_gss_fun:*
                23:  em_gss_mam:*
                24:  em_gss_mus:*
                25:  em_gss_other:*
                26:  em_gss_pro:*
                27:  em_gss_rod:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1092	39.9	669	13	BM013981	BM013981 603639619
	2	1068.5	39.0	1033	14	BM926413	BM926413 AGENCOURT
	3	1029	37.6	760	13	BM016204	BM016204 603642659
	4	918	33.5	545	9	AL040762	AL040762 DKFZp434C
	5	903	33.0	929	14	BQ882401	BQ882401 AGENCOURT
	6	856.5	31.3	948	13	BM471690	BM471690 AGENCOURT
	7	750	27.4	894	12	BF584668	BF584668 602098406
	8	744.5	27.2	872	11	AK020318	AK020318 Mus muscu
	9	689.5	25.2	2676	11	AK014319	AK014319 Mus muscu
	10	659	24.1	943	9	AL520322	AL520322 AL520322
	11	654	23.9	405	9	AA243820	AA243820 zr67g06.r
	12	650	23.7	550	13	BJ075336	BJ075336 BJ075336
	13	648	23.7	803	12	BG298756	BG298756 602396681
	14	637	23.3	780	12	BG587938	BG587938 EST489713
	15	625	22.8	559	12	BF692596	BF692596 602248949
c	16	621	22.7	606	14	BQ262340	BQ262340 fz85a08.x
	17	621	22.7	681	14	BU006064	BU006064 QGG9N16.y
c	18	604.5	22.1	932	17	AZ670821	AZ670821 ENTJN69TF
	19	602.5	22.0	913	17	BH155700	BH155700 ENTRO54TR
c	20	599	21.9	880	17	AZ687805	AZ687805 ENTLU52TF
	21	596.5	21.8	871	17	AZ682350	AZ682350 ENTKB16TF
c	22	590.5	21.6	891	17	AZ682250	AZ682250 ENTKT68TR
c	23	585.5	21.4	853	17	AZ679807	AZ679807 ENTHI67TR
	24	583	21.3	2564	11	AK020022	AK020022 Mus muscu
	25	582.5	21.3	787	12	BG584063	BG584063 EST485823
	26	581	21.2	586	13	BJ336461	BJ336461 BJ336461
c	27	579	21.1	661	12	BG302675	BG302675 f136b04.x
c	28	578	21.1	575	10	AW342644	AW342644 fi81f06.x
	29	577.5	21.1	886	17	AZ540627	AZ540627 ENTEQ18TF
c	30	577	21.1	835	13	BJ349604	BJ349604 BJ349604
	31	576.5	21.0	947	17	AZ683753	AZ683753 ENTIL96TF
	32	576	21.0	547	13	BJ333166	BJ333166 BJ333166
c	33	574.5	21.0	885	17	AZ686798	AZ686798 ENTFZ28TR
	34	574	21.0	542	13	BJ339421	BJ339421 BJ339421
c	35	572.5	20.9	860	17	AZ543293	AZ543293 ENTDQ80TF
	36	572.5	20.9	886	17	BH139685	BH139685 ENTNA47TR
	37	571	20.8	715	14	BQ869512	BQ869512 QGD6G09.y
c	38	570	20.8	814	13	BJ352304	BJ352304 BJ352304
c	39	570	20.8	820	13	BJ356540	BJ356540 BJ356540
	40	570	20.8	1019	12	BG248052	BG248052 602359987
	41	569	20.8	1341	11	AY108285	AY108285 Zea mays
	42	565	20.6	657	14	BQ802058	BQ802058 WHE2821_F
c	43	564	20.6	636	13	BJ386538	BJ386538 BJ386538
c	44	564	20.6	750	13	BJ445883	BJ445883 BJ445883
	45	562	20.5	1813	11	U66688	U66688 Homo sapien

ALIGNMENTS

```
RESULT 1
BM013981
LOCUS       BM013981                669 bp    mRNA    linear   EST 30-OCT-2001
DEFINITION  603639619F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5415556 5',
            mRNA sequence.
ACCESSION   BM013981
VERSION     BM013981.1  GI:16528335
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 669)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
```

```
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12060  row: f  column: 05
            High quality sequence stop: 664.
FEATURES             Location/Qualifiers
     source          1..669
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5415556"
                     /clone_lib="NIH_MGC_87"
                     /tissue_type="mammary adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.383 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      203 a    122 c    162 g    182 t
ORIGIN

Alignment Scores:
Pred. No.:              2.18e-123     Length:        669
Score:                  1092.00       Matches:       215
Percent Similarity:     99.54%        Conservative:  2
Best Local Similarity:  98.62%        Mismatches:    0
Query Match:            39.87%        Indels:        1
DB:                     13            Gaps:          0

US-09-873-409-7 (1-541) x BM013981 (1-669)

Qy     292 AlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPhe 311
           :::|||||||||:::||||||||||||||||||||||||||||||||||||||||||||||
Db      17 AGTGTTTTCTTTAATGTAATCCATAGCAGTTATTGCATTGGAGCAGCAGTCCCTCACTTT 76

Qy     312 GluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLys 331
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      77 GAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTTATTGATAAGAAA 136

Qy     332 ProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrVal 351
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     137 CCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTG 196

Qy     352 GluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGly 371
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     197 GAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGT 256

Qy     372 LeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGly 391
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     257 CTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGG 316

Qy     392 LysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMet 411
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     317 AAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATG 376

Qy     412 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 431
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     377 GTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTG 436

Qy     432 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 451
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     437 GTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGA 496

Qy     452 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 471
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     497 GATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTT 556

Qy     472 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 491
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     557 ATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATGAGT 616

Qy     492 GlyGlyGlnLysGlnArg-IleAlaIleAlaArgAlaLeuValArgAsnPro 508
           |||||||||||||||||| |||||||||||||||||||||||||||||||||
Db     617 GGAGGGCAGAAACAGAGGCATCGCAATTGCTCGTGCCTTAGTTCGAAACCCC 668

RESULT 2
BM926413
LOCUS       BM926413              1033 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6600788 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5764845
            5', mRNA sequence.
ACCESSION   BM926413
VERSION     BM926413.1  GI:19376792
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1033)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12818  row: o  column: 22
            High quality sequence start: 3
            High quality sequence stop: 686.
FEATURES             Location/Qualifiers
     source          1..1033
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5764845"
                     /clone_lib="NIH_MGC_114"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
                     male brains, age range 23-27 yo. Library is oligo-dT
                     primed and directionally cloned (EcoRV site is  destroyed
                     upon cloning).  Average insert size 1.5 kb, insert size
                     range 1-3 kb.  Library is normalized and enriched for
                     full-length  clones and was constructed by C. Gruber
                     (Invitrogen).  Research  Genetics tracking code 019. Note:
                     this is a NIH_MGC Library."
BASE COUNT      289 a    212 c    218 g    312 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              3.46e-120     Length:        1033
Score:                  1068.50       Matches:       229
Percent Similarity:     93.17%        Conservative:  3
Best Local Similarity:  91.97%        Mismatches:    17
Query Match:            39.01%        Indels:        4
DB:                     14            Gaps:          0

US-09-873-409-7 (1-541) x BM926413 (1-1033)

Qy     293 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 312
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     180 GTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAA 239

Qy     313 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 332
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     240 ACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTCATTGATAAGAAACCC 299
```

```
Qy     333 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 352
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     300 AGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGAA 359

Qy     353 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 372
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     360 TTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCTG 419

Qy     373 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 392
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     420 AATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAG 479

Qy     393 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 412
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     480 AGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGTG 539

Qy     413 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 432
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     540 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTGGTT 599

Qy     433 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 452
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     600 AGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGAGAT 659

Qy     453 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 472
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     660 GATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATC 719

Qy     473 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 492
           ||||||||||||||  |||||||||||||||||||  ||||||  |||||| ||||||||
Db     720 ATGGAGTTTCCTAATAA-TTTAATACATTGGTAGGGGAANAAGGAGCTCCAATGAGTGGA 778

Qy     493 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 512
               |||    ||||||  ||||||  ||||||||||||||||::: |||
Db     779 AGCCAGAACCG-AGGATCCCAATTGCTCCTGCCTTAGTTCGAAACCCCAGGAT-CTGAAT 836

Qy     513 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeu 532
           ||||||||||||||||||||||||||||||:::           |||  |||  |||
Db     837 TTAGATGAAGCTACGTCTGCCCTGGATTCAGAAACCAGTCAACTGGTCA--GCTGGCCTG 894

Qy     533 GluLysAspThrProArgTyrSerPhe 541
           ||||||:::|||||||||      |||
Db     895 GAGAAGGAAACCCCCCGGNAATTCTTT 921

RESULT 3
BM016204
LOCUS       BM016204                 760 bp    mRNA    linear   EST 30-OCT-2001
DEFINITION  603642659F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5418615 5',
            mRNA sequence.
ACCESSION   BM016204
VERSION     BM016204.1  GI:16530558
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 760)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12068  row: e  column: 16
            High quality sequence stop: 740.
FEATURES             Location/Qualifiers
     source          1..760
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5418615"
                     /clone_lib="NIH_MGC_87"
                     /tissue_type="mammary adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.383 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      229 a    150 c    187 g    193 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              1.52e-115       Length:         760
Score:                  1029.00         Matches:        209
Percent Similarity:     98.12%          Conservative:   0
Best Local Similarity:  98.12%          Mismatches:     2
Query Match:            37.57%          Indels:         2
DB:                     13              Gaps:           0

US-09-873-409-7 (1-541) x BM016204 (1-760)

Qy     330 LysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGly 349
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3 AAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGA 62

Qy     350 ThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeu 369
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63 ACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTG 122

Qy     370 LysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGly 389
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 AAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGC 182

Qy     390 SerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPhe 409
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 AGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTT 242

Qy     410 IleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle 429
           |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 ATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATN 302

Qy     430 GlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr 449
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 GGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTAT 362

Qy     450 GlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyr 469
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     363 GGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTAT 422

Qy     470 AspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGln 489
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     423 GATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAA 482

Qy     490 MetSerGlyGlyGlnLysGlnArgIleAlaIle-AlaArgAlaLeuValArgAsnProLy 509
           ||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db     483 ATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGGCTCGTGCCTTAGTTCGAAACCCCAA 542

Qy     509 sIleLeuIleLeu-AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValG 529
           ||||||||||    ||||||||||||||||||||||||||||||||||||||||||||||
Db     543 GATTCTGATTTTCAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTC 602

Qy     529 lnAlaAlaLeuGluLysAspThrProArgTyrSer 540
           |||||||||||||||||||||||||||||||||||
Db     603 AAGCTGCACTGGAGAAGGATACCCCCAGGTATTCC 637

RESULT 4
AL040762
LOCUS       AL040762                 545 bp    mRNA    linear   EST 29-FEB-2000
DEFINITION  DKFZp434C1815_r1 434 (synonym: htes3) Homo sapiens cDNA clone
```

```
             DKFZp434C1815 5', mRNA sequence.
ACCESSION    AL040762
VERSION      AL040762.1  GI:5409708
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 545)
  AUTHORS    Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
             Wiemann,S.
  TITLE      EST (Bloecker, et al.)
  JOURNAL    Unpublished (1999)
COMMENT      Contact: Bloecker H
             MIPS
             Am Klopferspitz 18a D-82152 Martinsried, Germany
             This is the 5' sequence of the clone insert
             Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
             Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
             sequenced by GBF (National Research Centre for Biotechnology Ltd.,
             Braunschweig/Germany) within the cDNA sequencing consortium of the
             German Genome Project.
             s1 sequence also available.
             This clone (DKFZp434C1815) is available at the RZPD in Berlin.
             Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
             Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1..545
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp434C1815"
                     /clone_lib="434 (synonym: htes3)"
                     /tissue_type="testis"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT      171 a     89 c    138 g    147 t
ORIGIN

Alignment Scores:
Pred. No.:               4e-102        Length:         545
Score:                   918.00        Matches:        181
Percent Similarity:      100.00%       Conservative:   0
Best Local Similarity:   100.00%       Mismatches:     0
Query Match:             33.52%        Indels:         0
DB:                      9             Gaps:           0

US-09-873-409-7 (1-541) x AL040762 (1-545)

Qy     333 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 352
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3 AGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGAA 62

Qy     353 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 372
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63 TTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCTG 122

Qy     373 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 392
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 AATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAG 182

Qy     393 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 412
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 AGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGTG 242

Qy     413 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 432
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTGGTT 302

Qy     433 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 452
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 AGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGAGAT 362

Qy     453 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 472
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     363 GATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATC 422

Qy     473 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 492
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     423 ATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGA 482

Qy     493 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 512
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     483 GGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATT 542

Qy     513 Leu 513
           |||
Db     543 TTA 545

RESULT 5
BQ882401
LOCUS       BQ882401                 929 bp    mRNA    linear   EST 16-AUG-2002
DEFINITION  AGENCOURT_8627902 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6291782
            5', mRNA sequence.
ACCESSION   BQ882401
VERSION     BQ882401.1  GI:22274409
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 929)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM2492  row: c  column: 15
            High quality sequence stop: 677.
FEATURES             Location/Qualifiers
     source          1..929
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:6291782"
                     /clone_lib="NIH_MGC_43"
                     /tissue_type="normal pigmented retinal epithelium"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Library constructed  by Ling Hong
                     in the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using ZAP-cDNA synthesis kit
                     (Stratagene) and  Superscript II RT (Life Technologies).
                     Note: this is a NIH_MGC Library. |"
BASE COUNT      281 a    176 c    237 g    234 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:               6.7e-100       Length:         929
Score:                   903.00         Matches:        188
Percent Similarity:      73.02%         Conservative:   61
Best Local Similarity:   55.13%         Mismatches:     58
Query Match:             32.97%         Indels:         34
DB:                      14             Gaps:           3

US-09-873-409-7 (1-541) x BQ882401 (1-929)

Qy     176 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThr-ValIl 195
```

```
            ||||||:::||||||||||||||||||||||||||||||::::::||||||||| |||||
Db        1 GCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGGTGAT 60

Qy      195 eAlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTy 215
            |||||||       |||:::||||||||:::
Db       61 TGCATTTGGAGGACAAAAGAAAGAACTTGAA----------------------------- 91

Qy      215 rAlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrTh 235
                                                             ||||||
Db       92 -----------------------------------------------------AGGTACAA 99

Qy      235 rGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSe 255
              :::||||||::::::||||||      ||||||||:::   |||   :::   :::||
Db      100 CAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTC 159

Qy      255 rLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyTh 275
            |:::||||||   :::       :::      :::|||  ||||||||||||||||||
Db      160 TATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGAC 219

Qy      275 rSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePh 295
            |:::|||:::|||:::||||||         |||:::||||||  ||||||  ||||||||
Db      220 CACCTTGGTCCTCTCAGGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTT 273

Qy      295 eSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAl 315
            |||||||:::       ::::::   :::|||   |||        |||    |||||
Db      274 TTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGC 333

Qy      315 aIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAs 335
            |   |||||||||||||||:::     ||||||::::::||||||   ||||||||||||||
Db      334 AAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGA 393

Qy      335 pAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAs 355
            |::::::|||    :::|||:::||||||::::::|||:::|||    :::||||||:::||
Db      394 CAGCTATTCGAAGAGTGGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAA 453

Qy      355 nValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuAr 375
            ||||    |||:::||||||||||||         :::||||||||||||||||||||||||::
Db      454 TGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAGATCTTGAAGGGTCTGAACCTGAA 513

Qy      375 gIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVa 395
            ::::::||||||:::||||||||||||||||||      :::|||   ||||||||||||
Db      514 GGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAAC 573

Qy      395 lValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAs 415
              |||||||||:::|||||||||||||||||||     :::|||    :::    ||||||
Db      574 AGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACA 633

Qy      415 nAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGl 435
              |||||||     :::|||||||||        |||:::    ||||||||||||||||||
Db      634 GGATATTAGGACCATAAATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGA 693

Qy      435 uProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValTh 455
            ||||||||||||    |||||||||:::    ||||||    |||||||||::::::||||||
Db      694 ACCTGTATTGTTTGCCACCACGATAGCTGAAAACATTCCCTATGGCCGTGAAAATGTCAC 753

Qy      455 rAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla-TyrAspPheIleMetGluP 475
            |    :::|||:::|||:::|||        :::||||||||||  |||||||||||||:::
Db      754 CATGGATGAGATTGAGAAAGCTGTCCAGGAAGCCAATGCCCTATGACTTTATCATGAAAC 813

Qy      475 heProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGly---G 494
               |||:::|||||:::|||||||||||         :::      |||:::|||||||||    |
Db      814 TGCCTCATAAATTTGACACCCTGGTTGGANAAAGAAGGGGCCAATTGAGTGGTGGGGCAC 873

Qy      494 lnLysGlnArgIleAlaIleAlaArg-AlaLeuVal-ArgAsnProLysIleLeu 511
            ||    :::          :::||||||  |||||||||    ||||||||||||||
Db      874 AAACCAAAGGATCTCCCTTGGCACGTGGCCCTGGTTCCCCAACCCCAAGATCCTC 928
```

RESULT 6
BM471690

```
LOCUS       BM471690                 948 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6465349 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5539117
            5', mRNA sequence.
ACCESSION   BM471690
VERSION     BM471690.1  GI:18520732
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 948)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12233  row: b  column: 14
            High quality sequence stop: 569.
FEATURES             Location/Qualifiers
     source          1. .948
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5539117"
                     /clone_lib="NIH_MGC_72"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.  Library constructed by Life
                     Technologies."
BASE COUNT      285 a    187 c    194 g    282 t
ORIGIN

Alignment Scores:
Pred. No.:              3.64e-94      Length:        948
Score:                  856.50        Matches:       186
Percent Similarity:     89.67%        Conservative:  5
Best Local Similarity:  87.32%        Mismatches:    14
Query Match:            31.27%        Indels:        8
DB:                     13            Gaps:          1

US-09-873-409-7 (1-541) x BM471690 (1-948)

Qy      293 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 312
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      205 GTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAA 264

Qy      313 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 332
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      265 ACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTTATTGATAAGAAACCC 324

Qy      333 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 352
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      325 AGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGAA 384

Qy      353 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 372
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      385 TTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCTG 444

Qy      373 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 392
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      445 AATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAG 504

Qy      393 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 412
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      505 AGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGTG 564

Qy      413 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle-GlyValVa 432
            |||||||||||||||||||||||||||||||||||||||||||||||||||| |||||||
Db      565 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTTGGAGTGGT 624
```

```
Qy    432 lSerGlnGluPro--ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr-GlyAr 451
Db    625 TAGTCAAGAGCCCTGGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGGACG 684

Qy    451 gAspAspValThrAspGluGlu-MetGluArgAlaAlaArgGluAla-AsnAlaTyrAsp 470
Db    685 AGATGATGGGACTGATGAACAAAATGGAAAAAGCAACAAGGGAAACCAAAGGCGAATGAA 744

Qy    471 PheIleMetGlu-PheProAsnLysPheAsnThrLeuVal---GlyGluLysGlyAlaGl 489
Db    745 TTTATCATGGAATTTCCCAAAAAAATTTAATACATTGGAAAGGGGAAAAAAGGACCTTCA 804

Qy    489 nMetSerGlyGlyGlnLysGlnArg 497
Db    805 AGGAATGGAAGGGCCAAAACCCAAG 829

RESULT 7
BF584668
LOCUS       BF584668                894 bp    mRNA    linear   EST 12-DEC-2000
DEFINITION  602098406F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4218385 5',
            mRNA sequence.
ACCESSION   BF584668
VERSION     BF584668.1  GI:11658386
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 894)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9798  row: 1  column: 02
            High quality sequence stop: 651.
FEATURES             Location/Qualifiers
     source          1..894
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4218385"
                     /clone_lib="NCI_CGAP_Co24"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.6 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      263 a    203 c    242 g    185 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              4.19e-81      Length:         894
Score:                  750.00        Matches:        156
Percent Similarity:     84.94%        Conservative:   47
Best Local Similarity:  65.27%        Mismatches:     36
Query Match:            27.38%        Indels:         4
DB:                     12            Gaps:           0

US-09-873-409-7 (1-541) x BF584668 (1-894)

Qy    293 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 312
Db      5 GTCTTCTTTTCCGTGTTAAT-GGAGCATTCAGTGTTGGACAGGCATCTCCAAATATTGAA 63

Qy    313 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 332
Db     64 GCCTTCGCCAATGCACGAGGAGCAGCTTATGAAGTCTTCAAAATAATTGATAATAAGCCC 123

Qy    333 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 352
Db    124 AGTATAGACAGCTTCTCAAAGAGTGGGCACAAACCAGACAACATACAAGGAAATCTGGAA 183

Qy    353 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 372
Db    184 TTTAAGAATATTCACTTCAGTTACCCATCTCGAAAAGAAGTTCAGATCTTGAAGGGCCTC 243

Qy    373 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 392
Db    244 AATCTGAAGGTGAAGAGCGGACAGACGGTGGCCCTGGTTGGCAACAGTGGCTGTGGAAAA 303

Qy    393 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 412
Db    304 AGCACAACTGTCCAGCTGATGCAAAGGCTCTACGACCCCCTAGATGGCATGGTCAGTATC 363

Qy    413 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 432
Db    364 GACGGACAGGACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCATTGGTGTGGTG 423

Qy    433 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 452
Db    424 AGTCAGGAACCTGTGCTGTTTGCCACCACGATCGCCGAGAACATTCGCTATGGCCGAGAA 483

Qy    453 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 472
Db    484 GATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTCATC 543

Qy    473 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 492
Db    544 ATGAAACTGCCCCACCAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGG 603

Qy    493 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 512
Db    604 GGACAGAAACAGAGAATCGC-ATTGCCCGGGCCCTGGTCCGCAATCCCAAGTGCCTTTC- 661

Qy    513 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 531
Db    662 GTGGACGAAGCCACCTCAGC-CTGGATACAGAAAGTGAAGCTGGGGTTCAGGCCGCA 717

RESULT 8
AK020318
LOCUS       AK020318                872 bp    mRNA    linear   HTC 19-JAN-2002
DEFINITION  Mus musculus adult male epididymis cDNA, RIKEN full-length enriched
            library, clone:9230106F14:similar to MULTIDRUG RESISTANCE PROTEIN 3
            (P-GLYCOPROTEIN 3) (MDR1A), full insert sequence.
ACCESSION   AK020318
VERSION     AK020318.1  GI:12860872
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (strain:C57BL/6J) adult male epididymis cDNA to mRNA,
            clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:9230106F14.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
```

```
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409 (6821), 685-690 (2001)
  MEDLINE   21085660
   PUBMED   11217851
REFERENCE   5  (bases 1 to 872)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C.,
            Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,
            Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F.,
            Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,
            Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,
            Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
            Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,
            Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and
            Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Gemomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. cDNA went
            through one round of normalization to Rot = 10.0 and subtraction to
            Rot = 185.2. Second strand cDNA was prepared with the primer
            adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cleaved
            with BamHI and XhoI. Vector: a modified pBluescript KS(+) after
            bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'
            end: BamHI. Host: DH10B.
FEATURES             Location/Qualifiers
     source          1. .872
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:9230106F14"
                     /db_xref="MGD:MGI:1911565"
                     /db_xref="taxon:10090"
                     /clone="9230106F14"
                     /sex="male"
                     /tissue_type="epididymis"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     CDS             86. .>872
                     /note="data source:SPTR, source key:P21447, evidence:ISS
                     putative
                     similar to MULTIDRUG RESISTANCE PROTEIN 3 (P-GLYCOPROTEIN
                     3) (MDR1A)"
                     /codon_start=1
                     /protein_id="BAB32070.1"
                     /db_xref="GI:12860873"
                     /db_xref="MGD:MGI:1924956"
                     /translation="MANSERTNGLQETNQRYGPLQEQVPKVGNQAVGPIEIFRFADNL
                     DIVLMTLGILASMINGATVPLMSLVLGEISDHLINGCLVQTNRTKYQNCSQTQEKLNE
                     DIIVLTLYYIGIGAAALIFGYVQISFWVITAARQTTRIRKQFFHSILAQDISWFDGSD
                     ICELNTRMTGDINKLCDGIGDKIPLMFQNISGFSIGLVISLIKSWKLSLVVLSTSPLI
                     MASSALCSRMIISLTSKELDAYSKAGAVAEEALSSIQTVTAFGA"
BASE COUNT      262 a    178 c    187 g    245 t
ORIGIN

Alignment Scores:
Pred. No.:              1.91e-80        Length:         872
Score:                  744.50          Matches:        151
Percent Similarity:     80.37%          Conservative:   21
Best Local Similarity:  70.56%          Mismatches:     27
Query Match:            27.18%          Indels:         15
DB:                     11              Gaps:           2

US-09-873-409-7 (1-541) x AK020318 (1-872)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            |||   ||||||||||||||||||::::::|||||||||   :::|||||||||   |||
Db      230 ATGACTCTGGGAATATTAGCATCCATGATAAATGGAGCCACCGTTCCTTTAATGTCCCTG 289

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
            ||||||||||||:::||||||:::||||||:::|||||||||||||||||||||
Db      290 GTTTTAGGAGAAATAAGTGATCATTTAATTAATGGATGCCTAGTACAAACTAACAGAACT 349

Qy       41 SerPhe----------------------------------------PheArgLeuThr 46
               :::                                             ||||||
Db      350 AAATATCAGAACTGTTCTCAGACTCAAGAAAAGCTGAATGAAGATATCATTGTGTTGACT 409

Qy       47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
            |||||||||:::|||||||||   |||||||||||||||||||||:::|||||||||
Db      410 CTATATTATATTGGAATAGGAGCAGCTGCCCTCATTTTTGGCTATGTACAGATTTCCTTC 469

Qy       67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
            |||:::|||||||||||||||||||||   |||||||||||||||||||||||||||:::
Db      470 TGGGTCATAACTGCAGCCCGGCAAACCACAAGAATCCGAAAACAGTTTTTCATTCAATT 529

Qy       87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
            |||||||||||||||   |||||||||      ||||||   ||||||||||||||||||
Db      530 TTGGCACAAGACATCAGCTGGTTTGATGGCAGTGACATCTGTGAACTTAACACCCGCATG 589

Qy      107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
            |||   ||||||:::|||:::   |||||||||||||||||||||   |||:::||||||
Db      590 ACTGGTGACATCAACAAACTCTGTGATGGTATTGGAGATAAGATCCCTCTGATGTTTCAG 649

Qy      126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
            |||:::|||   |||||||||||||||   :::   |||:::|||   |||||||||:::
```

```
Db        650 AACATATCTGGGTTTTCTATTGGCCTGGTGATAAGTTTGATAAAAAGCTGGAAACTCTCC 709

Qy        146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
              |||||   ||||||||||||||||||||||||||||||:::|||   ||||||||||||
Db        710 CTGGTGGTTCTGTCTACATCTCCTCTCATAATGGCTTCATCGGCACTGTGCTCTAGGATG 769

Qy        166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
              :::||||||||||||||||||||||||   |||||||||||||||||||||||||||||||
Db        770 ATTATCTCATTGACCAGCAAGGAGCTGGATGCCTATTCCAAAGCTGGGGCTGTGGCTGAA 829

Qy        186 GluValLeuSerSerIleArgThrValIleAlaPheArgAla 199
              |||   ||||||||||:::||||||   ||||||   |||
Db        830 GAAGCCTTGTCATCAATCCAAACAGTCACAGCCTTTGGAGCC 871

RESULT 9
AK014319
LOCUS       AK014319               2676 bp    mRNA    linear   HTC 19-JAN-2002
DEFINITION  Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length
            enriched library, clone:3222401P09:ATP-binding cassette, sub-family
            B (MDR/TAP), member 8, full insert sequence.
ACCESSION   AK014319
VERSION     AK014319.1  GI:12852089
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (strain:C57BL/6J) 14, 17 days embryo head cDNA to
            mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:3222401P09.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409 (6821), 685-690 (2001)
  MEDLINE   21085660
   PUBMED   11217851
REFERENCE   5  (bases 1 to 2676)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C.,
            Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,
            Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F.,
            Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,
            Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,
            Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
            Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,
            Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and
            Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Gemomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. Second
            strand cDNA was prepared with the primer adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cleaved
            with BamHI and XhoI. cDNA of size comprised between 3 and 7 kb was
            selected before cloning. Vector: a modified pBluescript KS(+) after
            bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'
            end: BamHI. Host: DH10B.
FEATURES             Location/Qualifiers
     source          1..2676
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:3222401P09"
                     /db_xref="MGD:MGI:1907488"
                     /db_xref="taxon:10090"
                     /clone="3222401P09"
                     /tissue_type="head"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="14, 17 days embryo"
     gene            1..2676
                     /gene="Abcb8"
     CDS             72..2225
                     /gene="Abcb8"
                     /note="ATP-binding cassette, sub-family B (MDR/TAP),
                     member 8
                     data source:MGD, source key:MGI:1351667, evidence:ISS
                     putative"
                     /codon_start=1
                     /protein_id="BAB29270.1"
                     /db_xref="GI:12852090"
                     /db_xref="MGD:MGI:1351667"
                     /translation="MLVHLFRFGIRGGPVPGWSLQSLRFQTFSAARSSDDRLSSHLLR
                     TVAQLRVQLRAHLPRAPPASHWSPSAWCWVGGTLVVPAVLWQHPRLCLIALCEAKESP
                     PAQPTRAPELRFNWKLFWHFLHPHLLALGAAIVLALGAALVNVQIPLLLGQLVEIVAK
                     YTRDHMGSFVSESRKLSVQLLLLYGVQGLLTFGYLVLLSHIGERMAMDMRKALFSSLL
```

```
                    RQDIAFFDAKKTGQLVSRLTTDVQEFKSSFKLVISQGLRSCTQVIGSLVSLSMLSPRL
                    TLMLAVVTPALMGVGTLMGSGLRKLSRQCQEQIARATGVADEALGNVRTVRAFAMEKR
                    EEERYQAELESCCCKAEELGRGIALFQGLSNIAFNCMVLGTLFIGGSLVAGQQLKGGD
                    LMSFLVASQTVQRSMASLSVLFGQVVRGLSAGARVFEYMALSPVIPLTGGYCIPNKDI
                    RGSITFQNVTFSYPCRPGFNVLKDFTLKLPSGKIVALVGQSGGGKTTVASLLERFYDP
                    EAGSVTLDGHDLRTLNPSWLRGQVIGFISQEPVLFATTIMENIRFGKLDASDEEVYTA
                    AREANAHEFISSFPDGYSTVVGERGTTLSGGQKQRLAIARALIKQPTVLILDEATSAL
                    DAESERVVQEALDRASAGRTVLVIAHRLSTVRAAHSIIVMANGQVCEAGTHEELLKKG
                    GLYSELIRRQTLDASLTSTPPAEKPEDPKSCQSKA"
     polyA_signal   2653..2658
                    /gene="Abcb8"
                    /note="putative"
     polyA_site     2676
                    /gene="Abcb8"
                    /note="putative"
BASE COUNT      544 a    785 c    756 g    591 t
ORIGIN

Alignment Scores:
Pred. No.:              7e-73          Length:         2676
Score:                  689.50         Matches:        195
Percent Similarity:     52.42%         Conservative:   98
Best Local Similarity:  34.88%         Mismatches:     187
Query Match:            25.17%         Indels:         79
DB:                     11             Gaps:           20

US-09-873-409-7 (1-541) x AK014319 (1-2676)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
              ::::::|||   :::    |||:::|||||||||               :::||||||
Db      471 ATTGTGCTGGCCTTAGGTGCGGCACTAGTGAAT------------GTGCAGATCCCCTTG 518

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIleSer------------------GlyCysLeu 34
              :::|||||||::::::   :::    :::                     |||
Db      519 CTCCTGGGCCAGCTGGTGGAGATTGTCGCCAAGTACACGAGGGACCACATGGGGAGTTTC 578

Qy       35 ValGlnThrAsnThrTyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyVal 54
              |||        :::                :::|||  |||    |||       ||||||
Db      579 GTGTCTGAGTCCCGTAAGCTCAGCGTCCAGCTGCTCCTACTGTAC---------GGTGTT 629

Qy       55 AlaAlaLeuIle---PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArg 73
                    |||:::    |||||||||:::    :::         |||          |||
Db      630 CAGGGACTGCTGACCTTTGGATACCTAGTGCTGCTGTCCCACATT------GGTGAGCGC 683

Qy       74 GlnThrLysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrp 93
                       :::|||||||      |||    |||:::|||    ||||||||     :::
Db      684 ATGGCCATGGACATGCGGAAAGCCCTTTTCAGCTCCCTGCTCCGGCAAGACATTGCTTTC 743

Qy       94 PheAspSerCysAspIleGlyGluLeuAsnThrArgMet---ThrAspIleAspLys--- 111
              ||||||:::            |||:::|||   :::|||:::    ||||||:::  :::
Db      744 TTTGATGCCAAAAAGACAGGGCAGCTAGTGAGTCGCTTGACTACTGATGTGCAAGAATTC 803

Qy      112 --------------------IleSerAspGlyIleGlyAspLysIleAlaLeuLeuPhe 124
                                  ||||||   |||:::
Db      804 AAGTCATCCTTCAAGCTTGTCATCTCCCAGGGACTG------------------------ 839

Qy      125 GlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu---ValLysGlyTrpLys 143
                         |||     ||||||      |||   |||     :::          :::
Db      840 ---CGCAGCTGCACCCAGGTGATTGGTAGCCTGGTGTCCCTGTCTATGCTGTCCCCTCGC 896

Qy      144 LeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSer 163
              |||||||||:::                :::|||     :::|||
Db      897 CTTACCCTGATGCTGGCTGTCGTCACACCCGCCCTCATGGGAGTGGGCACCCTGATGGGC 956

Qy      164 ArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaVal 183
                    :::      |||:::    :::                  ::::::|||      |||
Db      957 TCAGGCCTCCGAAAGCTGTCTCGCCAGTGTCAGGAGCAGATTGCCAGGGCAACAGGTGTA 1016

Qy      184 AlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGlu 203
              |||:::|||    |||       ::::::||||||||    ||||||      :::::::::|||
Db     1017 GCAGATGAGGCCCTTGGCAATGTTCGGACTGTGCGGGCCTTCGCGATGGAGAAGAGGGAG 1076

Qy      204 LeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrp 223
                :::|||                                      :::
Db     1077 GAGGAACGC---------------------------------------TATCAAGCAGAA 1097

Qy      224 LeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAsp 243
              |||   ||||||                                      |||::::::
Db     1098 CTGGAGTCATGCTGCTGT------------------------------AAAGCAGAAGAA 1127

Qy      244 PheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhe------ 261
              |||      |||  ||||||           |||        :::   |||
Db     1128 CTGGGC------AGGGGCATCGCCTTGTTCCAAGGCTCTCCAACATCGCTTTCAACTGT 1181

Qy      262 PheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGly 281
                 :::   ||||||               :::   |||    ||||||:::
Db     1182 ATGGTCTTGGGCACC------------CTGTTCATTGGGGGCTCCCTTGTGGCTGGACAG 1229

Qy      282 Glu---ProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
              :::      |||     :::   :::   |||                :::|||     |||
Db     1230 CAGCTGAAAGGGGGAGACCTCATGTCCTTCCTGGTGGCTTCCCAGACAGTACAGAGGTCT 1289

Qy      301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGly--- 319
                 :::                                   :::   ||||||
Db     1290 ATGGCCAGCCTCTCTGTCCTGTTTGGTCAG---------------GTGGTACGTGGGCTG 1334

Qy      320 ---AlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSer 338
                 |||          :::|||:::     :::          |||     |||
Db     1335 AGTGCGGGAGCCCGAGTCTTCGAATACATGGCCCTGAGCCCTGTCATCCCA------TTG 1388

Qy      339 ThrAlaGlyTyr------LysProGluSerIleGluGlyThrValGluPheLysAsnVal 356
              |||  ||||||               :::     |||    |||::::::    |||:::||||||
Db     1389 ACCGGGGGCTACTGCATCCCCAACAAGGACATTCGTGGTTCCATCACCTTCCAAAATGTC 1448

Qy      357 SerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIle 376
              :::|||:::||||||    ||||||              :::||||||         |||::::::
Db     1449 ACCTTCAGCTACCCCTGCAGACCTGGCTTCAATGTGCTCAAGGACTTCACCCTGAAGCTG 1508

Qy      377 LysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValVal 396
                 ||||||:::    ||||||||||||||      :::|||    ||||||:::||||||
Db     1509 CCCTCTGGCAAGATTGTGGCTCTTGTGGGCCAGTCTGGGGGAGGAAAGACCACAGTTGCC 1568

Qy      397 GlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAsp 416
                 ||||||:::|||   |||||||||:::   |||   :::   :::|||     :::|||
Db     1569 TCCTTGCTGGAACGCTTCTATGACCCTGAAGCTGGCTCGGTGACGTTGGATGGGCATGAC 1628

Qy      417 IleArgAlaLeuAsnValArgHisTyrArgAspHis---IleGlyValValSerGlnGlu 435
              :::|||   ||||||          |||          ||||||     :::|||||||||
Db     1629 CTGCGAACTCTCAACCCCTCCTGGCTCCGGGGCCAGGTCATAGGTTTCATCAGCCAGGAG 1688

Qy      436 ProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThr 455
              ||||||||||||   |||||||||      ||||||::::::|||:::    |||     :::
Db     1689 CCAGTCCTGTTTGCAACAACCATCATGGAGAATATCCGATTTGGGAAGCTGGATGCTTCC 1748

Qy      456 AspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPhe 475
              |||||||||:::       ||||||||||||||||||||:::::::||||||       |||
Db     1749 GATGAAGAGGTGTACACAGCTGCACGAGAAGCCAATGCCCACGAGTTCATCAGCAGCTTC 1808

Qy      476 ProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLys 495
              |||:::   ::::::|||:::|||||||||:::|||          :::|||||||||||||||
Db     1809 CCCGATGGCTACAGCACTGTGGTTGGTGAGCGGGGCACAACCTTGTCTGGTGGCCAGAAG 1868

Qy      496 GlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGlu 515
              ||||||:::||||||||||||||||||::::::    |||     :::|||||||||||||||
Db     1869 CAGCGCCTAGCCATCGCACGTGCCCTCATCAAGCAGCCCACAGTGCTGATCCTGGACGAG 1928

Qy      516 AlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
              ||||||||||||||||||:::||||||:::          ||||||    ||||||::::::
Db     1929 GCCACCAGTGCGCTAGATGCAGAATCCGAGAGGGTGGTACAGGAGGCCCTGGACCGG 1985

RESULT 10
AL520322
LOCUS       AL520322                943 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL520322 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB006YC15 5
            prime, mRNA sequence.
ACCESSION   AL520322
```

```
VERSION     AL520322.1  GI:12783815
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 943)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1. .943
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DB006YC15"
                     /clone_lib="LTI_NFL004_NBC2"
                     /sex="male"
                     /tissue_type="neuroblastoma cells"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector.  Library was normalized. Library was constructed
                     by Life Technologies.  Contact : Feng Liang Life
                     Technologies, a division of Invitrogen 9800 Medical Center
                     Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
                     8371 Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      253 a    225 c    253 g    211 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:                 7.19e-70      Length:         943
Score:                     659.00        Matches:        131
Percent Similarity:        72.36%        Conservative:   47
Best Local Similarity:     53.25%        Mismatches:     66
Query Match:               24.06%        Indels:         2
DB:                        9             Gaps:           1

US-09-873-409-7 (1-541) x AL520322 (1-943)

Qy    291 LeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHis 310
          |||   :::|||      :::|||:::        :::       :::|||
Db      2 CTGTTATTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTT 61

Qy    311 PheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLys 330
                :::|||   |||:::   :::|||   ||||||      :::|||:::|||
Db     62 GCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATTGAAAAA 121

Qy    331 LysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThr 350
             |||   ||||||::::::||||||   |||      |||   ::::::||||||
Db    122 ACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAAT 181

Qy    351 ValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLys 370
          |||   |||      |||   ||||||||||||:::||||||   |||   :::|||:::
Db    182 GTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAG 241

Qy    371 GlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySer 390
          ||||||:::|||      ||||||:::|||:::||||||||||||      :::|||
Db    242 GGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGT 301

Qy    391 GlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIle 410
          |||||||||||||||||||||||||:::|||   |||||||||      |||   :::
Db    302 GGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGGCAGGGAAAGTG 361

Qy    411 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 430
          ::::::|||      :::|||:::   |||||||||:::      |||   |||:::|||
Db    362 CTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGC 421

Qy    431 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 450
          :::|||||||||||||||:::||||||      :::|||:::    ||||||   ||||||
Db    422 ATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGA 481

Qy    451 ArgAspAsp------ValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla 468
             :::          |||:::   ||||||:::   |||||||||:::|||||||||
Db    482 GACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATA 541

Qy    469 TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAla 488
          :::   ||||||            ||||||||||::::::|||   ||||||:::||||||
Db    542 CATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACT 601

Qy    489 GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro 508
          |||:::||||||||||||||||||||||||||||||||||||||||||||||||   |||
Db    602 CAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCT 661

Qy    509 LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 528
             ||||||:::||||||||||||||||||||||||||||:::||||||:::      |||
Db    662 CATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTC 721

Qy    529 GlnAlaAlaLeuGluLys 534
          |||   ||||||:::|||
Db    722 CAAGAAGCCCTGGACAAA 739

RESULT 11
AA243820
LOCUS       AA243820              405 bp    mRNA    linear   EST 06-AUG-1997
DEFINITION  zr67g06.r1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:668506 5'
            similar to SW:MDR1_MOUSE P06795 MULTIDRUG RESISTANCE PROTEIN 1 ;,
            mRNA sequence.
ACCESSION   AA243820
VERSION     AA243820.1  GI:1874631
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 405)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.
            , Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
            ,T., Waterston,R. and Wilson,R.
  TITLE     WashU-Merck EST Project 1997
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1572   Std Error: 0.00
            Seq primer: -28m13 rev2 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1. .405
                     /organism="Homo sapiens"
                     /db_xref="GDB:5562473"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:668506"
                     /clone_lib="Soares_NhHMPu_S1"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
```

```
                  5,000 clones made from the same 3 libraries. The pools
                  consisted of I.M.A.G.E. clones 260232-265223,
                  340488-345479, and 484488-489479."
BASE COUNT      115 a     81 c     85 g    124 t
ORIGIN

Alignment Scores:
Pred. No.:                7.44e-70       Length:         405
Score:                    654.00         Matches:        128
Percent Similarity:       100.00%        Conservative:   1
Best Local Similarity:    99.22%         Mismatches:     0
Query Match:              23.88%         Indels:         0
DB:                       9              Gaps:           0

US-09-873-409-7 (1-541) x AA243820 (1-405)

Qy    291 LeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHis 310
          |||:::|||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     18 CTCTCTGTTTCTTTAGTGTAATCCATAGCAGTTATTGCATTGGAGCAGCAGTCCCTCAC 77

Qy    311 PheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLys 330
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     78 TTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTTATTGATAAG 137

Qy    331 LysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThr 350
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    138 AAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACT 197

Qy    351 ValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLys 370
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    198 GTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAA 257

Qy    371 GlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySer 390
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    258 GGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGT 317

Qy    391 GlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIle 410
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    318 GGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATC 377

Qy    411 MetValAspGluAsnAspIleArgAla 419
          |||||||||||||||||||||||||||
Db    378 ATGGTGGATGAGAATGACATCAGAGCT 404

RESULT 12
BJ075336
LOCUS       BJ075336                550 bp    mRNA    linear   EST 11-DEC-2001
DEFINITION  BJ075336 NIBB Mochii normalized Xenopus tailbud library Xenopus
            laevis cDNA clone XL057d10 5', mRNA sequence.
ACCESSION   BJ075336
VERSION     BJ075336.1  GI:17520252
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 550)
  AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
            ,Y.
  TITLE     Expressed genes in X. laevis embryo
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1..550
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="XL057d10"
                     /clone_lib="NIBB Mochii normalized Xenopus tailbud
                     library"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 25"
                     /note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
                     were oligo-dT primed and directionally cloned. Staging
                     according to  Nieuwkoop and Faber. Library is subtracted
                     and was constructed by N. Garrett  and A.M. Zorn,
                     (Wellcome/CRC Institute). "
BASE COUNT      162 a    109 c    144 g    135 t
ORIGIN

Alignment Scores:
Pred. No.:                3.81e-69       Length:         550
Score:                    650.00         Matches:        121
Percent Similarity:       85.79%         Conservative:   36
Best Local Similarity:    66.12%         Mismatches:     26
Query Match:              23.73%         Indels:         0
DB:                       13             Gaps:           0

US-09-873-409-7 (1-541) x BJ075336 (1-550)

Qy    349 GlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIle 368
          |||    :::||||||||   :::    |||    ||||||   ||||||   ::::::|||
Db      2 GGCGATATAGAATTCAAAGGCATTCATTTTGCCTACCCATGCCGTCCTGGTGTGCAAATT 61

Qy    369 LeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsn 388
          |||   ||||||||||||:::::::::::|||:::|||:::||||||||||||::::::
Db     62 CTCTCTGGACTGAACCTGAAAGTGGAAGCTGGTAAAACCATCGCTTTGGTTGGTATGAGT 121

Qy    389 GlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGly 408
          |||   ||||||||||||   :::|||||||||||||||   |||||||||      |||
Db    122 GGTTGTGGTAAAAGCACTACCATACAGTTGCTGCAGAGGTTCTATGATCCATCACATGGA 181

Qy    409 PheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHis 428
             :::   :::|||   :::|||||||||:::||||||||||:::      |||::::::
Db    182 GAGGTTACTTTGGATGGACACGATATCCGTTCACTAAATGTAAAATGGTTGAGGGAAAAT 241

Qy    429 IleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLys 448
          |||||||||||||||||||||||||||||||||||||||||||:::   ||||||:::
Db    242 ATTGGCGTAGTGAGTCAGGAGCCAGTGCTCTTTGGCACAACCATTGCAGAAAATATCCGC 301

Qy    449 TyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla 468
          ||||||   :::   |||||||||   |||:::|||:::||||||:::||||||||||||
Db    302 TACGGAGGGGAATCAGTAACAGACAACGAAATTGAACAAGCGGCAAAAGAAGCCAATGCC 361

Qy    469 TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAla 488
          :::|||||||||         |||:::||||||   |||:::|||||||||:::||||||
Db    362 TTTGATTTTATATCCAGGCTTCCTGATAAATTTAAGACGATGGTGGGAGAAAGAGGTGCG 421

Qy    489 GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro 508
          |||:::||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    422 CAGTTGAGTGGGGGTCAGAAGCAGCGGATAGCCATTGCACGCGCCCTGGTACGGAACCCT 481

Qy    509 LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 528
          |||||||||:::||||||||||||||||||||||||||||||:::::::|||::::::   |||
Db    482 AAGATTCTACTTCTTGATGAAGCCACATCAGCTCTGGATACACAAAGTGAAGCTATTGTA 541

Qy    529 GlnAlaAla 531
          |||||||||
Db    542 CAAGCTGCC 550

RESULT 13
BG298756
LOCUS       BG298756                803 bp    mRNA    linear   EST 21-FEB-2001
DEFINITION  602396681F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4511157 5',
            mRNA sequence.
ACCESSION   BG298756
VERSION     BG298756.1  GI:13063728
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 803)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: The Cepko Laboratory
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM10394  row: b  column: 22
            High quality sequence stop: 698.
FEATURES             Location/Qualifiers
     source          1. .803
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4511157"
                     /clone_lib="NIH_MGC_94"
                     /tissue_type="retina"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 3.3 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      240 a    162 c    205 g    196 t
ORIGIN

Alignment Scores:
Pred. No.:                 1.25e-68       Length:         803
Score:                     648.00         Matches:        149
Percent Similarity:        68.99%         Conservative:   49
Best Local Similarity:     51.92%         Mismatches:     55
Query Match:               23.66%         Indels:         37
DB:                        12             Gaps:           3

US-09-873-409-7 (1-541) x BG298756 (1-803)

Qy     168 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 187
           |||   |||   |||||||||   |||||:::||||||||||||||||||||||||||||
Db      15 TCATTTACTGATAAGGAACTCCATGCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTC 74

Qy     188 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSer 207
           |||::::::||||||||||||||||         |||:::||||||||:::
Db      75 TTAGCAGCCATCAGAACTGTGATTGCGTTGGA-GGACAAAAGAAGGAACTTGAA------ 127

Qy     208 PheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCys 227

Db     127 ------------------------------------------------------------ 127

Qy     228 ValLeu***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 247
                         ||||||      ||||||::::::||||||      ||||||||
Db     128 ---------------AGGTACAATAACAACTTGGAAGAAGCTAAAAGGCTGGGGATAAAG 172

Qy     248 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 267
           :::   |||   :::   :::|||:::||||||   :::         :::      :::|||
Db     173 AAAGCTATCACGGCCAACATCTCCATGGGTGCAGCTTTTCTCCTTATCTATGCATCATAT 232

Qy     268 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 287
              |||||||||||||||||||||||||||:::::::::   |||       |||:::
Db     233 GCTCTGGCATTCTGGTATGGGACTTCCTTGGTCATCTCCAAAGAA------TACTCTAT- 285

Qy     288 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 307
           |||   ||||||   ||||||||||||||:::        ::::::   :::|||   |||
Db     286 GGACAAGTGCTCACTGTCTTCTTTTCCGTGTTAATTGGAGCATTCAGTGTTGGACAGGCA 345

Qy     308 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 327
              |||:::   |||   ||||||   ||||||||||||||:::   :::|||::::::

Db     346 TCTCCAAATATTGAAGCCTTCGCCAATGCACGAGGAGCAGCTTATGAAGTCTTCAAAATA 405

Qy     328 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 347
           ||||||   ||||||||||||||:::||||||   :::|||:::||||||::::::|||
Db     406 ATTGATAATAAGCCCAGTATAGACAGCTTCTCAAAGAGTGGGCACAAACCAGACAACATA 465

Qy     348 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 367
           :::|||   :::||||||||||||:::   |||:::||||||||||||      ::::::
Db     466 CAAGGAAATCTGGAATTTAAGAATATTCACTTCAGTTACCCATCTCGAAAAGAAGTTCAG 525

Qy     368 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 387
           ||||||||||||||||||||||||::::::||||||||:::||||||||||||||||||
Db     526 ATCTTGAAGGGCCTCAATCTGAAGGTGAAGAGCGGACAGACGGTGGCCCTGGTTGGCAAC 585

Qy     388 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 407
           :::|||   ||||||||||||   ||||||||:::||||||||||||||||||   |||
Db     586 AGTGGCTGTGGAAAAAGCACAACTGTCCAGCTGATGCAAAGGCTCTACGACCCCTA-GAT 644

Qy     408 GlyPheIleMetValAspGluAsnAspIleArgAlaLeu-AsnValArgHis-TyrArg- 426
           |||   :::   :::|||      ||||||||   :::   ||||||||:::   :::|||
Db     645 GGCATGGTCAGTATCGACGGACAGGACATCAGAACCATCCAATGTGAGGTATCTGGAGGG 704

Qy     427 AspHisIleGlyValValSerGln-GluProValLeuPheGlyThrThrIleSerAsnAs 446
           ||||||   |||||||||||||||| ||||||   ||||||
Db     705 GATCAT---GGGGTGGTGAGTCAAGGAACCTGGCTTGTTTGCCAACCCAGATCGCGAGAA 761

Qy     446 nIleLysTyrGly 450
            |||:::||||||
Db     762 AATTCGTTATGGG 774

RESULT 14
BG587938
LOCUS       BG587938                 780 bp    mRNA    linear   EST 11-APR-2001
DEFINITION  EST489713 MHAM Medicago truncatula/Glomus versiforme mixed EST
            library cDNA clone pMHAM-51L19 5' end, mRNA sequence.
ACCESSION   BG587938
VERSION     BG587938.1  GI:13603002
KEYWORDS    EST.
SOURCE      Medicago truncatula/Glomus versiforme mixed EST library.
  ORGANISM  Medicago truncatula/Glomus versiforme mixed EST library
            Eukaryota; mixed EST libraries.
REFERENCE   1  (bases 1 to 780)
  AUTHORS   Harrison,M.J., Liu,J., Town,C.D., Van Aken,S., Utterback,T., Cho,J.
            and Fraser,C.M.
  TITLE     ESTs from roots of Medicago truncatula after colonization with
            Glomus versiforme, 2001
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Harrison M.J.
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73401
            Tel: 580-223-5810
            Fax: 580-221-7380
            Email: mjharrison@noble.org
            Noble EST name: N387011e TIGR sequence name: MTDDM70TK More
            information is available at: http://www.medicago.org
            Seq primer: SKmod (CTA gAA CTA gtg gAT CC).
FEATURES             Location/Qualifiers
     source          1. .780
                     /organism="Medicago truncatula/Glomus versiforme mixed EST
                     library"
                     /cultivar="Medicago truncatula genotype A17"
                     /db_xref="taxon:119092"
                     /clone="pMHAM-51L19"
                     /clone_lib="MHAM"
                     /tissue_type="roots colonized with Glomus versiforme"
                     /dev_stage="Roots harvested at 10, 17, 22, 31 and 38 days
                     post-inoculation with Glomus versiforme. The library was
                     made from a mixture of RNA from each of these stages."
                     /lab_host="E. coli strain XLOLR"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; cDNA was prepared from polyA+ enriched RNA from
                     roots harvested at 10, 17, 22, 31 and 38 days
```

```
                 post-inoculation with Glomus versiforme. The cDNA was
                 directionally ligated into the Unizap XR vector from
                 Stratagene and packaged using Gigapack III Gold packaging
                 extracts. Plasmids containing cDNA inserts were excised
                 from the recombinant lambda-Zap phage using Ex-assist
                 helper phage and propagated in XLOLR cells."
BASE COUNT      234 a    156 c    184 g    206 t
ORIGIN

Alignment Scores:
Pred. No.:              2.68e-67        Length:         780
Score:                  637.00          Matches:        126
Percent Similarity:     71.43%          Conservative:   39
Best Local Similarity:  54.55%          Mismatches:     66
Query Match:            23.26%          Indels:         0
DB:                     12              Gaps:           0

US-09-873-409-7 (1-541) x BG587938 (1-780)

Qy    304 IleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHis 323
           :::|||  :::  |||               |||    ||||||  ||||||
Db      3 TTGGGGCAATCTGCACCGAGCATGGCTGCATTTACAAAGGCTAGAGTTGCGGCTGCTAAG 62

Qy    324 IlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLys 343
          ||||||::::::||||||    :::|||    ||||||      |||   :::|||   :::
Db     63 ATTTTCCGGATAATTGATCACCAGCCTGGTATAGATAGAAACAGTGAATCTGGATTGGAA 122

Qy    344 ProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArg 363
             |||:::::: |||   ||||||   |||||||||   |||:::||||||||||||
Db    123 TTAGAGACAGTTACTGGACTTGTTGAACTGAAAAATGTGGACTTCTCTTATCCATCAAGA 182

Qy    364 ProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAla 383
          |||   :::  ||||||          :::|||   :::   :::|||:::|||:::|||
Db    183 CCTGAAGTTCTGATCCTCAATGATTTCTCCTTGAGTGTTCCTGCCGGAAAGACCATAGCT 242

Qy    384 LeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyr 403
          |||||||||   :::||||||||||||||||||||||||   |||::::::|||   |||
Db    243 TTAGTTGGTAGCAGCGGCTCTGGCAAGAGCACTGTTGTTTCTCTTATTGAGAGATTCTAT 302

Qy    404 AspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArg 423
          ||||||      |||     :::|||:::|||    :::||||||:::   |||   ::::::
Db    303 GATCCAACTTCAGGACAAGTAATGTTGGATGGGCATGACATTAAAACTTTGAAGCTTAAA 362

Qy    424 HisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIle 443
                |||      ||||||:::||||||||||||||||   ||||||   |||||||||
Db    363 TGGTTGAGGCAACAAATAGGACTAGTGAGCCAAGAACCTGCTTTGTTTGCCACCACGATT 422

Qy    444 SerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAla 463
                ||||||       ||||||  |||             |||:::|||   ||||||
Db    423 CGAGAAAATATACTCTTGGGAAGGCCTGATGCAAACCAGGTTGAGATTGAAGAAGCTGCT 482

Qy    464 ArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuVal 483
          |||    |||||||||:::    ||||||::::::   |||       |||   |||   |||
Db    483 AGGGTTGCTAATGCTCATTCATTCATCATCAAACTTCCTGAAGGCTTTGAAACTCAGGTA 542

Qy    484 GlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAla 503
          ||||||:::|||   |||:::||||||||||||||||||||||||||||||||||||
Db    543 GGAGAAAGAGGACTACAACTTTCTGGAGGACAAAAACAAAGAATAGCAATAGCAAGGGCA 602

Qy    504 LeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGlu 523
          :::::::::||||||   ||||||:::||||||||||||||||||||||||||||||
Db    603 ATGCTAAAAAATCCAGCAATTCTTCTCCTCGATGAGGCAACAAGTGCATTGGACTCTGAA 662

Qy    524 SerLysSerAlaValGlnAlaAlaLeuGluLys 534
          |||:::      ||||||   ||||||::::::
Db    663 TCAGAAAAGCTGGTGCAAGAAGCACTTGACCGG 695

RESULT 15
BF692596
LOCUS       BF692596                 559 bp    mRNA    linear   EST 22-DEC-2000
DEFINITION  602248949F1 NIH_MGC_62 Homo sapiens cDNA clone IMAGE:4334100 5',
            mRNA sequence.
ACCESSION   BF692596
VERSION     BF692596.1  GI:11978004
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 559)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1203  row: a  column: 13
            High quality sequence stop: 555.
FEATURES             Location/Qualifiers
     source          1..559
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4334100"
                     /clone_lib="NIH_MGC_62"
                     /tissue_type="melanotic melanoma, high MDR"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: skin; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc);
                     Double-stranded cDNA was prepared from cell line RNA.  5'
                     and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).   Average
                     insert size 1.75 kb (range 0.9-4.0 kb).  15/15 colonies
                     contained  inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA)."
BASE COUNT      161 a    118 c    141 g    139 t
ORIGIN

Alignment Scores:
Pred. No.:              4.66e-66        Length:         559
Score:                  625.00          Matches:        134
Percent Similarity:     96.43%          Conservative:   1
Best Local Similarity:  95.71%          Mismatches:     4
Query Match:            22.82%          Indels:         3
DB:                     12              Gaps:           0

US-09-873-409-7 (1-541) x BF692596 (1-559)

Qy    403 TyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVal 422
          |||||||||||||||||||||||||   |||||||||||||||||||||||||||||
Db      1 TATGATCCGGATGATGGCTTTATCACGGTGGATGAGAATGACATCAGAGCTTTAAATGTG 60

Qy    423 ArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThr 442
          ||||||||||||||||||   |||||||||||||||||||||   ||||||||||||||
Db     61 CGGCATTATCGAGACCATAT-GGAGTGGTTAGTCAAGAGCCTGT-TTGTTCGGGACCACC 118

Qy    443 IleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAla 462
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    119 ATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCA 178

Qy    463 AlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeu 482
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    179 GCAAGGGAAGCAAATGCATATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTG 238

Qy    483 ValGlyGluLys-GlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaAr 502
          ||||||||||   |||||||||||||||||||||||||||||||||||||||||||||
Db    239 GTAGGGGAAAACAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCG 298
```

```
Qy     502 gAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSe 522
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     299 TGCCTTAGTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTC 358

Qy     522 rGluSerLysSerAlaValGlnAlaAlaLeuGluLysAspThrProArgTyrSerPhe 541
           ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||
Db     359 AGAAAGCGAGTCAGCTGTTCAAGCTGCACTGGAGAAGGATACCCCCAGGTATTCATTT 416
```

Search completed: March 31, 2003, 13:51:37
Job time : 3734.22 secs

```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 31, 2003, 02:24:20 ; Search time 79.2267 Seconds
                                           (without alignments)
                                           2094.145 Million cell updates/sec

Title:          US-09-873-409-7
Perfect score:  2739
Sequence:       1 MILGILASLVNGACLPLMPL..........SESKSAVQAALEKDTPRYSF 541

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters:      882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115422_19259/app_query.fasta_1.7544
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_578_@runat_27032003_115422_19259 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Issued_Patents_NA:*
                1:  /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
                2:  /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
                3:  /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
                4:  /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
                5:  /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
                6:  /cgn2_6/ptodata/2/ina/backfiles1.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1552.5	56.7	4233	3	US-09-120-513-1	Sequence 1, Appli
	2	1552.5	56.7	4233	4	US-09-450-105-1	Sequence 1, Appli
	3	1552.5	56.7	4669	2	US-08-752-447-1	Sequence 1, Appli
	4	1552.5	56.7	4669	4	US-09-316-167-1	Sequence 1, Appli
	5	1546.5	56.5	4646	1	US-08-181-471-2	Sequence 2, Appli
	6	1546.5	56.5	6505	2	US-08-793-610-5	Sequence 5, Appli
	7	1546.5	56.5	9318	2	US-08-793-610-6	Sequence 6, Appli
	8	1545	56.4	4264	2	US-08-784-649A-1	Sequence 1, Appli
	9	1545	56.4	4264	2	US-08-784-649A-5	Sequence 5, Appli
	10	1544.5	56.4	4669	6	5206352-3	Patent No. 5206352
	11	1525.5	55.7	4669	2	US-08-583-276-18	Sequence 18, Appl
	12	960	35.0	4224	1	US-08-612-521-1	Sequence 1, Appli
	13	935.5	34.2	4047	2	US-08-612-734B-1	Sequence 1, Appli
	14	916	33.4	4800	2	US-08-612-734B-3	Sequence 3, Appli
	15	914	33.4	4002	2	US-08-996-545-1	Sequence 1, Appli
	16	914	33.4	4002	2	US-08-996-545-3	Sequence 3, Appli
	17	914	33.4	4002	4	US-09-328-320-1	Sequence 1, Appli
	18	914	33.4	4002	4	US-09-328-320-3	Sequence 3, Appli
	19	898	32.8	2726	1	US-08-461-823-1	Sequence 1, Appli
	20	897.5	32.8	3924	1	US-08-395-246C-1	Sequence 1, Appli
	21	849.5	31.0	6143	1	US-08-612-521-3	Sequence 3, Appli
	22	785.5	28.7	3792	4	US-09-351-224E-10	Sequence 10, Appl
	23	747.5	27.3	3924	2	US-08-996-644-3	Sequence 3, Appli
	24	747.5	27.3	3924	3	US-09-352-552-3	Sequence 3, Appli
	25	747.5	27.3	3927	2	US-08-996-644-1	Sequence 1, Appli
	26	747.5	27.3	3927	3	US-09-352-552-1	Sequence 1, Appli
	27	688.5	25.1	3999	4	US-09-351-224E-9	Sequence 9, Appli
	28	646	23.6	2376	1	US-08-394-880B-1	Sequence 1, Appli
	29	626.5	22.9	2061	4	US-09-061-764A-17	Sequence 17, Appl
	30	626	22.9	3909	1	US-08-232-537-1	Sequence 1, Appli
	31	621.5	22.7	1959	4	US-09-061-764A-4	Sequence 4, Appli
	32	574.5	21.0	2244	4	US-09-061-764A-18	Sequence 18, Appl
	33	571.5	20.9	1749	4	US-09-134-001C-1893	Sequence 1893, Ap
	34	529	19.3	1743	4	US-09-134-001C-2774	Sequence 2774, Ap
	35	519.5	19.0	7760	4	US-08-961-527-63	Sequence 63, Appl
	36	501.5	18.3	5120	3	US-08-772-270A-6	Sequence 6, Appli
	37	501.5	18.3	8370	2	US-08-488-706-1	Sequence 1, Appli
	38	491	17.9	4411529	4	US-09-103-840A-1	Sequence 1, Appli
	39	489	17.9	13188	4	US-08-961-527-70	Sequence 70, Appl
	40	489	17.9	4403765	4	US-09-103-840A-2	Sequence 2, Appli
	41	483	17.6	7721	3	US-08-772-270A-14	Sequence 14, Appl
c	42	482	17.6	28804	2	US-08-592-874-1	Sequence 1, Appli
c	43	482	17.6	28804	3	US-09-096-942-2	Sequence 2, Appli
c	44	482	17.6	28804	3	US-09-096-867-2	Sequence 2, Appli
	45	460	16.8	6492	4	US-08-961-527-188	Sequence 188, App

ALIGNMENTS

```
RESULT 1
US-09-120-513-1
; Sequence 1, Application US/09120513
; Patent No. 6025160
;  GENERAL INFORMATION:
;    APPLICANT:  Brun, Kimberly
;    APPLICANT:  Chenery, Richard
;    APPLICANT:  Ellens, Harma
;    APPLICANT:  Field, John
;    APPLICANT:  Yue, Lin
;    TITLE OF INVENTION:  POLYNUCLEOTIDE AND POLYPEPTIDE
;    TITLE OF INVENTION:  SEQUENCES ENCODING RAT MDR1B2 AND
;    TITLE OF INVENTION:  SCREENING METHODS THEREOF
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road
;      CITY:  King of Prussia
;      STATE:  PA
;      COUNTRY:
;      ZIP:  19406
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/120,513
;      FILING DATE:  22-JUL-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  King, William T
;      REGISTRATION NUMBER:  30,954
;      REFERENCE/DOCKET NUMBER:  GP50008
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-5015
```

```
;      TELEFAX:  610-270-5090
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4233 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
US-09-120-513-1

Alignment Scores:
Pred. No.:              1.69e-172      Length:          4233
Score:                  1552.50        Matches:         308
Percent Similarity:     72.20%         Conservative:    92
Best Local Similarity:  55.60%         Mismatches:      105
Query Match:            56.68%         Indels:          49
DB:                     3              Gaps:            5

US-09-873-409-7 (1-541) x US-09-120-513-1 (1-4233)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    170 ATGGCTCTGGGAACTCTCGCTGCTATCATCCACGGAACCCTGCTTCCCCTCCTGATGCTG 229

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
Db    230 GTGTTCGGATACATGACAGATAGTTTACCCAAGCAGAGACCCGCATTCTGCCGAGCGTT 289

Qy     31 ------SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu--------- 45
Db    290 ACTAATCAAAGTGAAATCAACAGTACACAGACCGTCAGCGACAGCAGTCTGGAGGAGGAC 349

Qy     46 ------------ThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    350 ATGGCCATGTACGCCTACTATTACACGGGCATTGGTGCCGGTGTGCTCATCGTTGCCTAC 409

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    410 ATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAAG 469

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    470 TTTTTCCATGCCATCATGAATCAGGAGATAGGCTGGTTTGACGTGAATGACGCTGGGGAG 529

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    530 CTCAACACCCGGCTCACAGATGACGTCTCCAAAATTAATGACGGAATTGGTGACAAACTT 589

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    590 GGAATGTTCTTTCAGTCCATAACGACATTTTCAGCCGGTTTTATAATAGGATTTATAAGT 649

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    650 GGTTGGAAGCTAACCCTTGTAATTTTGGCCGTCAGCCCTCTTATTGGGTTGTCATCTGCC 709

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    710 ATGTGGGCAAAGGTACTGACTTCATTTACTAATAAGGAACTCCAGGCTTATGCGAAAGCT 769

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    770 GGAGCAGTTGCCGAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCGTTTGGAGGACAA 829

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db    830 AAGAAGGAACTTGAA--------------------------------------------- 844

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db    845 ------------------------------------AGGTACAATAAAAATTTAGAAGAA 868

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db    869 GCTAAAAGAGTTGGCATAAAGAAAGCCATCACGGCCAACATTTCCATAGGTATTGCCTAC 928

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db    929 CTGTTGGTCTATGCGTCTTATGCACTGGCATTCTGGTATGGACCTCCTTGGTCCTCTCA 988

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db    989 AATGAA------TATTCTATTGGACAAGTGCTTACCGTCTTCTTCTCTATTTATTGGGG 1042

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db   1043 ACTTTCAGTATTGGACATTTAGCCCCAAACATAGAAGCCTTTGCAAATGCAAGAGGGGCA 1102

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   1103 GCCTATGAAATCTTCAAGATAATTGATAATGAGCCAAGCATTGACAGCTTCTCAACCAAG 1162

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   1163 GGACACAAACCAGACAGTATAATGGGAAATTTGGAATTTAAAAATGTTTACTTCAACTAC 1222

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   1223 CCATCACGAAGTGAAGTTAAGATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAG 1282

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   1283 ACGGTAGCCCTGGTTGGCAACAGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAG 1342

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   1343 AGGCTCTACGACCCCATAGAGGGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATC 1402

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   1403 AATGTGAGGTATCTGCGGGAAATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCC 1462

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   1463 ACCACGATTGCCGAAAACATTCGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAG 1522

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   1523 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTCATCATGAAACTGCCCCACAAATTTAAC 1582

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   1583 ACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATT 1642

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   1643 GCCCGGGCCCTGGTCCGCAACCCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTG 1702

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db   1703 GACACAGAAAGCGAAGCCGTGGTTCAGGCCGCTCTGGATAAG 1744

RESULT 2
US-09-450-105-1
; Sequence 1, Application US/09450105
; Patent No. 6169166
; GENERAL INFORMATION:
;  APPLICANT: Kimberly Anne Brun
;  APPLICANT: Richard James Chenery
;  APPLICANT: Harma Ellens
;  APPLICANT: John Anthony Feild
;  APPLICANT: Lin Yue
;  TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
;  TITLE OF INVENTION: ENCODING RAT MDR1B2 AND SCREENING METHODS THEREOF
;  FILE REFERENCE: GP-50008-D1
```

```
;  CURRENT APPLICATION NUMBER: US/09/450,105
;  CURRENT FILING DATE: 1999-11-29
;  EARLIER APPLICATION NUMBER: 09/120,513
;  EARLIER FILING DATE: 1998-07-22
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 4233
;   TYPE: DNA
;   ORGANISM: HOMO SAPIENS
US-09-450-105-1

Alignment Scores:
Pred. No.:              1.69e-172     Length:        4233
Score:                  1552.50       Matches:       308
Percent Similarity:     72.20%        Conservative:  92
Best Local Similarity:  55.60%        Mismatches:    105
Query Match:            56.68%        Indels:        49
DB:                     4             Gaps:          5

US-09-873-409-7 (1-541) x US-09-450-105-1 (1-4233)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    170 ATGGCTCTGGGAACTCTCGCTGCTATCATCCACGGAACCCTGCTTCCCCTCCTGATGCTG 229

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
Db    230 GTGTTCGGATACATGACAGATAGTTTTACCCAAGCAGAGACCCGCATTCTGCCGAGCGTT 289

Qy     31 ------SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu--------- 45
Db    290 ACTAATCAAAGTGAAATCAACAGTACACAGACCGTCAGCGACAGCAGTCTGGAGGAGGAC 349

Qy     46 ------------ThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    350 ATGGCCATGTACGCCTACTATTACACGGGCATTGGTGCCGGTGTGCTCATCGTTGCCTAC 409

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    410 ATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAAG 469

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    470 TTTTTCCATGCCATCATGAATCAGGAGATAGGCTGGTTTGACGTGAATGACGCTGGGGAG 529

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    530 CTCAACACCCGGCTCACAGATGACGTCTCCAAAATTAATGACGGAATTGGTGACAAACTT 589

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    590 GGAATGTTCTTTCAGTCCATAACGACATTTTCAGCCGGTTTTATAATAGGATTTATAAGT 649

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    650 GGTTGGAAGCTAACCCTTGTAATTTTGGCCGTCAGCCCTCTTATTGGGTTGTCATCTGCC 709

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    710 ATGTGGGCAAAGGTACTGACTTCATTTACTAATAAGGAACTCCAGGCTTATGCGAAAGCT 769

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    770 GGAGCAGTTGCCGAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCGTTTGGAGGACAA 829

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db    830 AAGAAGGAACTTGAA--------------------------------------------- 844

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db    845 ------------------------------------AGGTACAATAAAAATTTAGAAGAA 868

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db    869 GCTAAAAGAGTTGGCATAAAGAAAGCCATCACGGCCAACATTTCCATAGGTATTGCCTAC 928

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db    929 CTGTTGGTCTATGCGTCTTATGCACTGGCATTCTGGTATGGGACCTCCTTGGTCCTCTCA 988

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db    989 AATGAA------TATTCTATTGGACAAGTGCTTACCGTCTTCTTCTCTATTTTATTGGGG 1042

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db   1043 ACTTTCAGTATTGGACATTTAGCCCCAAACATAGAAGCCTTTGCAAATGCAAGAGGGGCA 1102

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   1103 GCCTATGAAATCTTCAAGATAATTGATAATGAGCCAAGCATTGACAGCTTCTCAACCAAG 1162

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   1163 GGACACAAACCAGACAGTATAATGGGAAATTTGGAATTTAAAAATGTTTACTTCAACTAC 1222

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   1223 CCATCACGAAGTGAAGTTAAGATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAG 1282

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   1283 ACGGTAGCCCTGGTTGGCAACAGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAG 1342

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   1343 AGGCTCTACGACCCCATAGAGGGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATC 1402

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   1403 AATGTGAGGTATCTGCGGGAAATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCC 1462

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   1463 ACCACGATTGCCGAAAACATTCGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAG 1522

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   1523 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTCATCATGAAACTGCCCCACAAATTTAAC 1582

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   1583 ACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATT 1642

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   1643 GCCCGGGCCCTGGTCCGCAACCCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTG 1702

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db   1703 GACACAGAAAGCGAAGCCGTGGTTCAGGCCGCTCTGGATAAG 1744

RESULT 3
US-08-752-447-1
; Sequence 1, Application US/08752447
; Patent No. 5994088
;  GENERAL INFORMATION:
;    APPLICANT:  Mechetner, Eugene
;    APPLICANT:  Roninson, Igor B
;    TITLE OF INVENTION:  Methods and Reagents for Preparing and
;    TITLE OF INVENTION:  Using Immunoligcal Agents Specific for P-glycoprotein
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McDonnell Boehnen Hulbert & Berghoff Ltd.
```

```
;      STREET:  300 South Wacker Drive, Seventh Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/752,447
;      FILING DATE:  15-NOV-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5994088nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  95,1121
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-913-0001
;      TELEFAX:  312-913-9808
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4669 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  5'UTR
;      LOCATION:  1..424
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  425..4264
;    FEATURE:
;      NAME/KEY:  3'UTR
;      LOCATION:  4265..4669
US-08-752-447-1

Alignment Scores:
Pred. No.:               1.99e-172     Length:         4669
Score:                   1552.50       Matches:        305
Percent Similarity:      73.29%        Conservative:   101
Best Local Similarity:   55.05%        Mismatches:     99
Query Match:             56.68%        Indels:         49
DB:                      2             Gaps:           6

US-09-873-409-7 (1-541) x US-08-752-447-1 (1-4669)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            |||:::::|||  ||||||::::::::::::||||||  |||||||||||   |||
Db    575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy     21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
            |||  |||||||||:::|||                          :::|||:::|||
Db    635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                     ::::::      |||         |||||
Db    695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                |||          ||||||  ||||||||         |||:::      |||
Db    755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
            ||||||:::|||  |||   :::   |||  ||||||      :::||||||||||||
Db    815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
            ||||||||:::::::::   |||::::||||||||||||||     |||:::||||||
Db    875 TTTTTCATGCTATTATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
            ||||||||||||:::|||     |||:::   |||||::::::|||||||||||||||
Db    935 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAAATT 994

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
             :::   ||||||:::|||:::||||||     |||       |||||      :::
Db    995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
            ||||||||||||||||||   |||:::   ||||||:::::       |||||||||
Db   1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
                 ::::::::::::   |||   |||   |||||||||    ||||||:::||||||
Db   1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
            ||||||||||||||||||||||||::::::||||||||||||||||||      |||
Db   1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
            :::|||||||||:::
Db   1235 AAGAAAGAACTTGAA--------------------------------------------- 1249

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                               ||||||   :::||||||::::::
Db   1250 -----------------------------------AGGTACAACAAAAATTTAGAAGAA 1273

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
            ||||||      |||||||||:::   |||     :::   :::|||:::||||||   :::
Db   1274 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 1333

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
                 :::      :::|||   :::||||||||||||||||||:::|||:::|||:::
Db   1334 CTGCTGATCTATGCATCTTATGCTGTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 1393

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
            ||||||          |||:::||||||      ||||||   ||||||||||||||:::
Db   1394 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1447

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
            ::::::   :::|||   |||   |||         |||   ||||||   ||||||||||||
Db   1448 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1507

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
            |||:::   ||||||::::::||||||   |||||||||||||||::::::|||     :::
Db   1508 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1567

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
            |||:::||||||::::::|||:::|||      :::||||||:::||||||      |||:::|||
Db   1568 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1627

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
            |||||||||         :::||||||||||||||||||:::::::::||||||:::
Db   1628 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1687

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
            ||||||||||||||||||      :::|||   ||||||||||||   |||||||||:::|||
Db   1688 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1747

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
            ||||||||||||||||      :::|||   :::   ||||||      |||||||||   :::
Db   1748 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1807

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
            |||||||||         |||:::   |||||||||||||||||||||||||||||||||
Db   1808 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1867

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
            |||||||||:::      ||||||:::||||||||||::::::||||||      :::|||:::|||
Db   1868 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1927

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
```

```
                :::|||   :::|||||||||||||||||||||||||:::    |||:::||||||:::
Db    1928 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1987

Qy     481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
           ||||||||||||||:::||||||||:::||||||||||||||||||||||||||||||||
Db    1988 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 2047

Qy     501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
           ||||||||||||||||||||||||||||||||||:::|||||||||||||||||||||||
Db    2048 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 2107

Qy     521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
           |||:::||||||::::::  ||||||   ||||||:::|||
Db    2108 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 2149

RESULT 4
US-09-316-167-1
; Sequence 1, Application US/09316167
; Patent No. 6365357
;  GENERAL INFORMATION:
;    APPLICANT:  Mechetner, Eugene
;    APPLICANT:  Roninson, Igor B
;    TITLE OF INVENTION:  Methods and Reagents for Preparing and
;    TITLE OF INVENTION:  Using Immunoligcal Agents Specific for P-glycoprotein
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McDonnell Boehnen Hulbert & Berghoff Ltd.
;      STREET:  300 South Wacker Drive, Seventh Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/316,167
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/752,447
;      FILING DATE:  15-NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 6365357nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  95,1121
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-913-0001
;      TELEFAX:  312-913-9808
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4669 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  5'UTR
;      LOCATION:  1..424
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  425..4264
;    FEATURE:
;      NAME/KEY:  3'UTR
;      LOCATION:  4265..4669
US-09-316-167-1

Alignment Scores:
Pred. No.:              1.99e-172     Length:         4669
Score:                  1552.50       Matches:        305
Percent Similarity:     73.29%        Conservative:   101
```

```
Best Local Similarity:  55.05%        Mismatches:     99
Query Match:            56.68%        Indels:         49
DB:                     4             Gaps:           6

US-09-873-409-7 (1-541) x US-09-316-167-1 (1-4669)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||::::::|||   ||||||::::::::::::||||||   |||||||||||   |||
Db     575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy      21 ValLeuGlyGluMetSerAsp-------------------------AsnLeuIleSer 31
           |||   |||||||||:::|||                         :::|||:::|||
Db     635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe----------------- 43
                 ::::::      |||         ||||||
Db     695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                 |||         ||||||   |||||||||         |||:::      |||
Db     755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
           ||||||:::|||   |||   :::   |||   ||||||       :::||||||||||||
Db     815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
           |||||||||:::::::::      |||:::|||||||||||||||     |||:::||||||
Db     875 TTTTTTCATGCTATTATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
           ||||||||||||:::|||   |||:::   ||||||::::::|||||||||||||||||||
Db     935 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAAATT 994

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
              :::   ||||||:::|||:::||||||      |||      ||||||      :::
Db     995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
           |||||||||||||||||||||   |||:::   ||||||::::::      |||||||||
Db    1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
              ::::::::::::   |||   |||      |||      ||||||:::||||||
Db    1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
           ||||||||||||||||||||||||::::::||||||||||||||||||||      |||
Db    1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy     201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
           :::|||||||||:::
Db    1235 AAGAAAGAACTTGAA--------------------------------------------- 1249

Qy     221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                            ||||||   :::||||||::::::
Db    1250 ------------------------------------AGGTACAACAAAAATTTAGAAGAA 1273

Qy     241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
           ||||||      |||||||||:::   |||      :::   :::|||:::||||||   :::
Db    1274 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 1333

Qy     261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
              :::      :::|||   :::||||||||||||||||||||:::|||:::|||:::
Db    1334 CTGCTGATCTATGCATCTTATGCTGTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 1393

Qy     281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
           ||||||      |||:::||||||   ||||||   ||||||||||||||||:::
Db    1394 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1447

Qy     301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
           ::::::   :::|||   |||   |||         |||   ||||||   ||||||||||||
Db    1448 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1507
```

```
Qy     321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
           |||:::   ||||||:::::||||||   ||||||||||||||:::::::|||   :::
Db    1508 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1567

Qy     341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
           |||:::||||||::::::|||:::|||   :::||||||:::||||||   |||:::|||
Db    1568 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1627

Qy     361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
           |||||||||         :::||||||||||||||||||||||||:::::::::||||||:::
Db    1628 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1687

Qy     381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
           ||||||||||||||||||   :::|||   ||||||||||||   |||||||||:::|||
Db    1688 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1747

Qy     401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
           ||||||||||||||||||   :::|||   :::   ||||||      |||||||||   :::
Db    1748 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1807

Qy     421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
           |||||||||      |||:::   ||||||||||||||||||||||||||||||||||||
Db    1808 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1867

Qy     441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
           |||||||||:::   ||||||:::||||||||||:::::::||||||   :::|||:::|||
Db    1868 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1927

Qy     461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
           :::|||   :::||||||||||||||||||||||||||||:::   |||:::||||||:::
Db    1928 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1987

Qy     481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
           ||||||||||||||||:::||||||||||:::||||||||||||||||||||||||||||
Db    1988 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 2047

Qy     501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
           |||||||||||||||||||||||||||||||:::||||||||||||||||||||||||||||
Db    2048 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 2107

Qy     521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
           |||:::||||||::::::   ||||||   ||||||:::|||
Db    2108 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 2149
```

RESULT 5
US-08-181-471-2
; Sequence 2, Application US/08181471
; Patent No. 5641508
; GENERAL INFORMATION:
; APPLICANT: Li, Lingna
; APPLICANT: Lishko, Valeryi K.
; TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL
; TITLE OF INVENTION: COMPOSITIONS TO HAIR FOLLICLES
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Thomas Fitting
; STREET: 12526 High Bluff Drive, Suite 300
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92130
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/181,471
; FILING DATE: 13-JAN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/041,553
; FILING DATE: 02-APR-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: ANT0029P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-792-3680
; TELEFAX: 619-792-8477
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4646 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 425..4267
US-08-181-471-2

Alignment Scores:

Pred. No.:	1.01e-171	Length:	4646
Score:	1546.50	Matches:	305
Percent Similarity:	73.10%	Conservative:	100
Best Local Similarity:	55.05%	Mismatches:	100
Query Match:	56.46%	Indels:	49
DB:	1	Gaps:	6

US-09-873-409-7 (1-541) x US-08-181-471-2 (1-4646)

```
Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||::::::|||   ||||||:::::::::::::||||||   ||||||||||||   |||
Db     575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy      21 ValLeuGlyGluMetSerAsp-------------------------AsnLeuIleSer 31
           |||   ||||||||:::|||                          :::|||:::|||
Db     635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                       ::::::      |||   ||||||
Db     695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                 |||         ||||||   ||||||||||         |||:::      |||
Db     755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
           ||||||:::|||   |||   :::   |||   ||||||      :::||||||||||||
Db     815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
           ||||||||:::::::::   |||:::||||||||||||||||      |||:::||||||
Db     875 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
           ||||||||||||:::|||   |||:::   ||||||::::::   |||||||||||||||
Db     935 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 994

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
              :::   ||||||:::|||:::||||||      |||      ||||||      :::
Db     995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
           ||||||||||||||||||||   |||:::   ||||||::::::      |||||||||
Db    1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
              ::::::::::::   |||   |||   |||||||||   ||||||:::||||||
Db    1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
```

```
                ||||||||||||||||||||||:::::::||||||||||||||||||||        |||
Db      1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy       201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
             :::||||||||:::
Db      1235 AAGAAAGAACTTGAA--------------------------------------------- 1249

Qy       221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                               ||||||   :::||||||:::::::
Db      1250 -----------------------------------AGGTACAACAAAAATTTAGAAGAA 1273

Qy       241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
             ||||||     ||||||||:::   |||   :::   :::|||:::||||||     :::
Db      1274 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 1333

Qy       261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
                   :::      :::|||   |||||||||||||||||||||:::|||:::|||:::
Db      1334 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 1393

Qy       281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
             ||||||      |||:::||||||   ||||||   |||||||||||||||:::
Db      1394 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1447

Qy       301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
             ::::::   :::|||   |||   |||      |||   ||||||   ||||||||||||
Db      1448 GCTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1507

Qy       321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
             |||:::   ||||||::::::||||||   ||||||||||||||||:::::::|||   :::
Db      1508 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1567

Qy       341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
             |||:::||||||:::::::|||:::|||      :::||||||:::||||||   |||:::|||
Db      1568 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1627

Qy       361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
             |||||||||      :::||||||||||||||||||||||||||:::::::::||||||:::
Db      1628 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1687

Qy       381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
             ||||||||||||||||||      :::|||   ||||||||||||   |||||||||:::|||
Db      1688 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1747

Qy       401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
             |||||||||||||||   :::|||   :::   ||||||      |||||||||   :::
Db      1748 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1807

Qy       421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
             |||||||||      |||:::   ||||||||||||||||||||||||||||||||
Db      1808 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1867

Qy       441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
             |||||||||:::   ||||||:::||||||||:::::::||||||   :::|||:::|||
Db      1868 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1927

Qy       461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
             :::|||   :::|||||||||||||||||||||||||||:::   |||:::|||||||::: 
Db      1928 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1987

Qy       481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
             ||||||||||||||||:::||||||||:::|||||||||||||||||||||||||||||||
Db      1988 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 2047

Qy       501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
             ||||||||||||||||||||||||||||||:::||||||||||||||||||||||||||||
Db      2048 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 2107

Qy       521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
             |||:::||||||::::::   ||||||   ||||||:::|||
Db      2108 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 2149
```

RESULT 6
US-08-793-610-5

```
; Sequence 5, Application US/08793610
; Patent No. 5858744
;  GENERAL INFORMATION:
;    APPLICANT:  BAUM, Christopher
;    APPLICANT:  STOCKING-HARBERS, Carol
;    APPLICANT:  OSTERTAG, Wolfram
;    TITLE OF INVENTION:  RETROVIRAL VECTOR HYBRIDS AND THE USE THEROF
;    TITLE OF INVENTION:  FOR GENE TRANSFER
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nikaido, Marmelstein, Murray & Oram LLP
;      STREET:  655 Fifteenth Street N.W. Suite 330
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/793,610
;      FILING DATE:  07-MAR-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P 44 31 973.8
;      FILING DATE:  08-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 195 03 952.1
;      FILING DATE:  07-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/03175
;      FILING DATE:  10-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berman, Richard J.
;      REGISTRATION NUMBER:  39,105
;      REFERENCE/DOCKET NUMBER:  P1614-7007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)638-5000
;      TELEFAX:  (202)638-4810
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  6505 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA
US-08-793-610-5

Alignment Scores:
Pred. No.:              1.78e-171       Length:         6505
Score:                  1546.50         Matches:        305
Percent Similarity:     73.10%          Conservative:   100
Best Local Similarity:  55.05%          Mismatches:     100
Query Match:            56.46%          Indels:         49
DB:                     2               Gaps:           6

US-09-873-409-7 (1-541) x US-08-793-610-5 (1-6505)

Qy         1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
             |||:::::::|||   ||||||::::::::::::::||||||   ||||||||||||   |||
Db      1967 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 2026

Qy        21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
             |||   |||||||||:::|||                           :::|||:::|||
Db      2027 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 2086

Qy        32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                               ::::::   |||      ||||||
Db      2087 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 2146

Qy        44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                   |||      ||||||   |||||||||      |||:::      |||
```

```
Db   2147 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 2206

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db   2207 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 2266

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db   2267 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 2326

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db   2327 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 2386

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db   2387 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 2446

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db   2447 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 2506

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db   2507 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 2566

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db   2567 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 2626

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db   2627 AAGAAAGAACTTGAA--------------------------------------------- 2641

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db   2642 ------------------------------------AGGTACAACAAAAATTTAGAAGAA 2665

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db   2666 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 2725

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db   2726 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 2785

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db   2786 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 2839

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db   2840 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 2899

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   2900 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 2959

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   2960 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 3019

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   3020 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 3079

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   3080 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 3139

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   3140 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 3199

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   3200 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 3259

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   3260 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 3319

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   3320 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 3379

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   3380 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 3439

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   3440 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 3499

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db   3500 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 3541
```

RESULT 7
US-08-793-610-6
; Sequence 6, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
; APPLICANT: BAUM, Christopher
; APPLICANT: STOCKING-HARBERS, Carol
; APPLICANT: OSTERTAG, Wolfram
; TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEROF
; TITLE OF INVENTION: FOR GENE TRANSFER
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 Fifteenth Street N.W. Suite 330
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/793,610
; FILING DATE: 07-MAR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE P 44 31 973.8
; FILING DATE: 08-SEP-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 195 03 952.1
; FILING DATE: 07-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/03175
; FILING DATE: 10-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Berman, Richard J.
; REGISTRATION NUMBER: 39,105
; REFERENCE/DOCKET NUMBER: P1614-7007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)638-5000
; TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9318 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular

```
;    MOLECULE TYPE:  DNA
US-08-793-610-6

Alignment Scores:
Pred. No.:              3.25e-171      Length:          9318
Score:                  1546.50        Matches:         305
Percent Similarity:     73.10%         Conservative:    100
Best Local Similarity:  55.05%         Mismatches:      100
Query Match:            56.46%         Indels:          49
DB:                     2              Gaps:            6

US-09-873-409-7 (1-541) x US-08-793-610-6 (1-9318)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db   1926 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 1985

Qy     21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db   1986 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 2045

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db   2046 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 2105

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db   2106 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 2165

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db   2166 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 2225

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db   2226 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 2285

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db   2286 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 2345

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db   2346 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 2405

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db   2406 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 2465

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db   2466 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 2525

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db   2526 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 2585

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db   2586 AAGAAAGAACTTGAA--------------------------------------------- 2600

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db   2601 ------------------------------------AGGTACAACAAAAATTTAGAAGAA 2624

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db   2625 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 2684

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db   2685 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 2744

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db   2745 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 2798

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db   2799 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 2858

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   2859 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 2918

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   2919 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 2978

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   2979 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 3038

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   3039 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 3098

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   3099 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 3158

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   3159 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 3218

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   3219 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 3278

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   3279 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 3338

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   3339 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 3398

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   3399 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 3458

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db   3459 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 3500

RESULT 8
US-08-784-649A-1
; Sequence 1, Application US/08784649A
; Patent No. 5830697
;  GENERAL INFORMATION:
;    APPLICANT:  Sikic, Branimir I
;    APPLICANT:  Chen, Gang
;    TITLE OF INVENTION:  P-GLYCOPROTEIN MUTANT RESISTANT TO
;    TITLE OF INVENTION:  CYCLOSPORIN MODULATION
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  2200 Sand Hill Road
;      CITY:  Menlo Park
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
```

```
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/784,649A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sherwood, Pamela J
;      REGISTRATION NUMBER:  Reg.No. 5830697 36,677
;      REFERENCE/DOCKET NUMBER:  06037/007001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-322-5070
;      TELEFAX:  415-854-0875
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4264 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-784-649A-1

Alignment Scores:
Pred. No.:              1.31e-171      Length:          4264
Score:                  1545.00        Matches:         306
Percent Similarity:     73.29%         Conservative:    100
Best Local Similarity:  55.23%         Mismatches:      98
Query Match:            56.41%         Indels:          50
DB:                     2              Gaps:            7

US-09-873-409-7 (1-541) x US-08-784-649A-1 (1-4264)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    292 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 351

Qy     21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
Db    352 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 411

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db    412 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 471

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    472 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 531

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    532 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 591

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    592 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 651

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    652 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAAATT 711

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    712 GGA---ATGTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 768

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    769 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 828

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    829 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 888

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    889 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 948

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db    949 AAGAAAGAACTTGAA--------------------------------------------- 963

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db    964 ------------------------------------AGGTACAACAAAAATTTAGAAGAA 987

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db    988 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 1047

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db   1048 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 1107

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db   1108 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1161

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db   1162 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1221

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   1222 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1281

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   1282 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1341

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   1342 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1401

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   1402 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1461

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   1462 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1521

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   1522 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1581

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   1582 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1641

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
Db   1642 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1701

Qy    481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
Db   1702 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 1761

Qy    501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
Db   1762 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 1821

Qy    521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
Db   1822 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 1863

RESULT 9
US-08-784-649A-5
; Sequence 5, Application US/08784649A
```

```
; Patent No. 5830697
;  GENERAL INFORMATION:
;    APPLICANT:  Sikic, Branimir I
;    APPLICANT:  Chen, Gang
;    TITLE OF INVENTION:  P-GLYCOPROTEIN MUTANT RESISTANT TO
;    TITLE OF INVENTION:  CYCLOSPORIN MODULATION
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  2200 Sand Hill Road
;      CITY:  Menlo Park
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/784,649A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sherwood, Pamela J
;      REGISTRATION NUMBER:  Reg.No. 5830697 36,677
;      REFERENCE/DOCKET NUMBER:  06037/007001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-322-5070
;      TELEFAX:  415-854-0875
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4264 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-784-649A-5

Alignment Scores:
Pred. No.:              1.31e-171     Length:        4264
Score:                  1545.00       Matches:       306
Percent Similarity:     73.29%        Conservative:  100
Best Local Similarity:  55.23%        Mismatches:    98
Query Match:            56.41%        Indels:        50
DB:                     2             Gaps:          7

US-09-873-409-7 (1-541) x US-08-784-649A-5 (1-4264)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    292 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 351

Qy     21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
Db    352 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 411

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db    412 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 471

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    472 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 531

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    532 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 591

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    592 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 651

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    652 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAAATT 711

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    712 GGA---ATGTTCCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 768

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    769 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 828

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    829 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 888

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    889 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 948

Qy    201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
Db    949 AAGAAAGAACTTGAA--------------------------------------------- 963

Qy    221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
Db    964 ------------------------------------AGGTACAACAAAAATTTAGAAGAA 987

Qy    241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
Db    988 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 1047

Qy    261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
Db   1048 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 1107

Qy    281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
Db   1108 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1161

Qy    301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
Db   1162 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1221

Qy    321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
Db   1222 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1281

Qy    341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
Db   1282 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1341

Qy    361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
Db   1342 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1401

Qy    381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
Db   1402 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1461

Qy    401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
Db   1462 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1521

Qy    421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
Db   1522 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1581

Qy    441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
Db   1582 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1641

Qy    461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
```

```
                 :::|||   :::|||||||||||||||||||||||||:::    |||:::||||||:::
Db     1642 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1701

Qy      481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
            ||||||||||||||:::|||||||||:::|||||||||||||||||||||||||||||||
Db     1702 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 1761

Qy      501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
            ||||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||
Db     1762 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 1821

Qy      521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
            |||:::||||||::::::    ||||||    ||||||:::|||
Db     1822 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 1863
```

RESULT 10
5206352-3
;Patent No. 5206352
;   APPLICANT: Roninson, Igor B.;Pastan Ira H.;Gottesman,
;Michael M.
;   TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA
;SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
;   NUMBER OF SEQUENCES: 4
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/07/622,836
;   FILING DATE: 24-SEP-1990
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: 892,575
;   FILING DATE: 01-AUG-1986
;   APPLICATION NUMBER: 845,610
;   FILING DATE: 28-MAR-1986
;SEQ ID NO:3:
;   LENGTH: 4669
5206352-3

Alignment Scores:

Pred. No.:	1.75e-171	Length:	4669
Score:	1544.50	Matches:	305
Percent Similarity:	72.92%	Conservative:	99
Best Local Similarity:	55.05%	Mismatches:	101
Query Match:	56.39%	Indels:	49
DB:	6	Gaps:	6

US-09-873-409-7 (1-541) x 5206352-3 (1-4669)

```
Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            |||::::::|||    ||||||::::::::::::||||||    |||||||||||    |||
Db      575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy       21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
            |||    |||||||||:::|||                          :::|||:::|||
Db      635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy       32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                     ::::::      |||      ||||||
Db      695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy       44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                  |||          ||||||   |||||||||         |||:::      |||
Db      755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy       62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
            ||||||:::|||   |||    :::    |||   ||||||      :::||||||||||||
Db      815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy       82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
            ||||||||:::::::::    |||:::||||||||||||||||      |||:::||||||
Db      875 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy      102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
            ||||||||||||:::|||   |||:::      ||||||:::      |||||||||||||||
Db      935 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATCAAGTTATTGGTGACAAAATT 994

Qy      121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
               :::   ||||||:::|||:::||||||          |||      ||||||         :::
Db      995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy      141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
            ||||||||||||||||||    |||:::   ||||||::::::         |||||||||
Db     1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy      161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
                 ::::::::::::    |||    |||    |||||||||    ||||||:::||||||
Db     1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy      181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
            ||||||||||||||||||||||||::::::||||||||||||||||||         |||
Db     1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy      201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
            :::||||||||||:::
Db     1235 AAGAAAGAACTTGAA--------------------------------------------- 1249

Qy      221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                                    ||||||   :::||||||::::::
Db     1250 ------------------------------------AGGTACAACAAAAATTTAGAAGAA 1273

Qy      241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
            ||||||      |||||||||:::   |||   :::    :::|||:::||||||      :::
Db     1274 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 1333

Qy      261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
                 :::      :::|||   |||||||||||||||||||:::|||:::|||:::
Db     1334 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 1393

Qy      281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
            ||||||      |||:::||||||   ||||||   |||||||||||||||:::
Db     1394 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1447

Qy      301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
            ::::::         |||   |||   |||   ||||||      ||||||||||||
Db     1448 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1507

Qy      321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
            |||:::   ||||||::::::||||||   |||||||||||||||:::::::|||      :::
Db     1508 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1567

Qy      341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
            |||:::||||||::::::|||:::|||     :::||||||:::||||||     |||:::|||
Db     1568 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1627

Qy      361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
            ||||||||    :::|||||||||||||||||||||||||::::::::|||||||:::
Db     1628 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1687

Qy      381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
            ||||||||||||||||||   :::|||   ||||||||||||   ||||||||||:::|||
Db     1688 ACGGTGGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1747

Qy      401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
            ||||||||||||||||   :::|||      :::   ||||||      ||||||||      :::
Db     1748 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATA 1807

Qy      421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
            ||||||||      |||:::   ||||||||||||||||||||||||||||||
Db     1808 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1867

Qy      441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
            |||||||||:::   ||||||:::|||||||||::::::||||||      :::|||:::|||
Db     1868 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1927

Qy      461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
            :::|||     :::|||||||||||||||||||||||:::      |||:::||||||:::
Db     1928 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1987

Qy      481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
```

```
             ||||||||||||||:::|||||||||:::|||||||||||||||||||||||||||
Db    1988 ACCCTGGTTGGAGAGAGAGGGGCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 2047

Qy     501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
             |||||||||||||||||||||||||||||||||:::|||||||||||||||||||||||
Db    2048 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 2107

Qy     521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
             |||:::||||||::::::     ||||||     ||||||:::|||
Db    2108 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 2149

RESULT 11
US-08-583-276-18
; Sequence 18, Application US/08583276
; Patent No. 5837536
;  GENERAL INFORMATION:
;    APPLICANT:  McDonagh, Kevin T.
;    APPLICANT:  Nienhuis, Arthur
;    APPLICANT:  Tolstoshev, Paul
;    TITLE OF INVENTION:  IMPROVED EXPRESSION OF HUMAN
;    TITLE OF INVENTION:  MULTIDRUG RESISTANCE GENES AND IMPROVED
;    TITLE OF INVENTION:  SELECTION OF CELLS TRANSDUCED WITH SUCH GENES
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carella, Byrne, Bain, Gilfillan,
;      ADDRESSEE:  Cecchi & Stewart
;      STREET:  6 Becker Farm Road
;      CITY:  Roseland
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch diskette
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  DW4.V2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/583,276
;      FILING DATE:  05-JAN-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/332,444
;      FILING DATE:  31-OCT-1994
;      APPLICATION NUMBER:  07/887,712
;      FILING DATE:  22-MAY-1992
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4669 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  singular
;      TOPOLOGY:  linear
;    MOLECULE TYPE:
;      DESCRIPTION:  Genomic DNA
US-08-583-276-18

Alignment Scores:
Pred. No.:              3.06e-169       Length:         4669
Score:                  1525.50         Matches:        301
Percent Similarity:     72.56%          Conservative:   101
Best Local Similarity:  54.33%          Mismatches:     103
Query Match:            55.70%          Indels:         49
DB:                     2               Gaps:           6

US-09-873-409-7 (1-541) x US-08-583-276-18 (1-4669)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
             |||::::::|||    ||||||:::::::::::::||||||    |||||||||||    |||
Db     575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy      21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
             |||    |||||||||:::|||                           :::|||:::|||
Db     635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                            ::::::      |||      ||||||
Db     695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                   |||          ||||||    |||||||||          |||:::          |||
Db     755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
             ||||||:::|||    |||    :::    |||    ||||||          :::||||||||||||
Db     815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
             ||||||||:::::::::      |||:::|||||||||||||||          |||:::|||:::
Db     875 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGCAG 934

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
             ||||||||||||:::|||      |||:::    ||||||::::::    |||||||||||||||
Db     935 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAACTTATTGGTGACAAAATT 994

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
                 :::    ||||||:::|||:::||||||          |||          ||||||          :::
Db     995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
             |||||||||||||||||||    |||:::    ||||||::::::    |||||||||
Db    1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
                 ::::::::::::    |||    |||    |||||||||    ||||||:::||||||
Db    1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
             ||||||||||||||||||||||||::::::|||||||||||||||||||          |||
Db    1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy     201 GluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrpPheTyrPhe 220
             :::||||||||:::
Db    1235 AAGAAAGAACTTGAA--------------------------------------------- 1249

Qy     221 ProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsnLeuLysAsp 240
                                                   ||||||    :::||||||::::::
Db    1250 ------------------------------------AGGTACAACAAAAATTTAGAAGAA 1273

Qy     241 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 260
             ||||||    ||||||||:::    |||    :::    :::|||:::||||||    :::
Db    1274 GCTAAAAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTC 1333

Qy     261 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 280
                 :::    :::|||    ||||||||||||||||||||:::|||:::|||:::
Db    1334 CTGCTGATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCA 1393

Qy     281 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 300
             ||||||      |||:::||||||    ||||||    ||||||||||||:::
Db    1394 GGGGAA------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGG 1447

Qy     301 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 320
             ::::::    :::|||    |||    |||          |||    ||||||    ||||||||||||
Db    1448 GCTTTTAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCA 1507

Qy     321 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 340
             |||:::    ||||||::::::||||||    ||||||||||||||:::::::|||    :::
Db    1508 GCTTATGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGT 1567

Qy     341 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 360
             |||:::|||||||::::::|||:::|||    :::||||||:::||||||    |||:::|||
Db    1568 GGGCACAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTAC 1627

Qy     361 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 380
             ||||||||       :::|||||||||||||||||||||||:::::::::||||||:::
Db    1628 CCATCTCGAAAAGAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAG 1687

Qy     381 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 400
```

```
              |||||  |||||||||  :::|||  |||||||||||  ||||||||:::|||
Db    1688 ACGGTCCCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAG 1747

Qy     401 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 420
           ||||||||||||||   :::|||   :::   |||          ||||||||    :::
Db    1748 AGGCTCTATGACCCCACAGAGGGGATGGTCAGTGTTCATGGACAGGATATTAGGACCATA 1807

Qy     421 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 440
           ||||||||      |||:::  ||||||||||||||||||||||||||||||||||||
Db    1808 AATGTAAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCC 1867

Qy     441 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 460
           |||||||||:::   ||||||:::||||||||||::::::||||||    :::|||:::|||
Db    1868 ACCACGATAGCTGAAAACATTCGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAG 1927

Qy     461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
           :::|||   :::||||||||||||||||||||||||||||:::   |||:::||||||:::
Db    1928 AAAGCTGTCAAGGAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCATAAATTTGAC 1987

Qy     481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
           |||||||||||||||:::|||   |||:::|||||||||||||||||||||||||||||||
Db    1988 ACCCTGGTTGGAGAGAGAGGGCCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATT 2047

Qy     501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
           ||||||||||||||||||||||||:::||||||||||||||||||||||||||||||||||
Db    2048 GCACGTGCCCTGGTTCGCAACCCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTG 2107

Qy     521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
           |||:::||||||::::::   ||||||   ||||||:::|||
Db    2108 GACACAGAAAGCGAAGCAGTGGTTCAGGTGGCTCTGGATAAG 2149
```

RESULT 12
US-08-612-521-1
; Sequence 1, Application US/08612521
; Patent No. 5786463
; GENERAL INFORMATION:
; APPLICANT: Peery, Robert B
; APPLICANT: Skatrud, Paul L
; APPLICANT: Thornewell, Susan J
; TITLE OF INVENTION: MULTIPLE DRUG RESISTANCE GENE OF
; TITLE OF INVENTION: CRYPTOCOCCUS NEOFORMANS
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patent Division/AEH
; STREET: Lilly Corporate Center
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: USA
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,521
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hamilton, Amy E
; REGISTRATION NUMBER: 33,894
; REFERENCE/DOCKET NUMBER: X-9693
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 317-276-3169
; TELEFAX: 317-276-1294
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4224 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4224
US-08-612-521-1

```
Alignment Scores:
Pred. No.:              1.47e-102     Length:        4224
Score:                  960.00        Matches:       210
Percent Similarity:     56.89%        Conservative:  112
Best Local Similarity:  37.10%        Mismatches:    182
Query Match:            35.05%        Indels:        62
DB:                     1             Gaps:          8
```

US-09-873-409-7 (1-541) x US-08-612-521-1 (1-4224)

```
Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||:::||||||::::::   ::::::      |||:::|||   |||||||||   |||
Db     466 ATGGTCTTGGGCCTTGTGCTTGCCGTTGCAGCCGGATCTTGCCAACCTTTAATGACTTTA 525

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
           :::   |||   ::::::   :::      :::      ::::::      |||
Db     526 ATATTCGGTCGATTAACCACGTCTTTCACGAATTATGCTGTCATT---GCGAACCAAATA 582

Qy      41 SerPhePheArgLeuThr------------------------------------------ 46
           |||      ||||||
Db     583 TCCCAAGGGGGACTTACTCCCGAGACGTCTGCGGCTTTACAAGCAGCCAAAGATGATCTC 642

Qy      47 ------------------------LeuTyrTyrValGlyIleGlyValAlaAlaLeuIle 58
                                   ||||||   :::   ||||||:::         :::
Db     643 AAGACTCAATCTGGCCACAATGCCCTTTATCTGATGGCCATTGGCATTGGAATGTTTCTA 702

Qy      59 PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIle 78
              ::::::   :::   :::|||   :::|||               :::|||||||||
Db     703 GCGACTTGGCTTTACATGTTCATCTGGAATGTTACTGGCGAGCTCAACTCTAAAAGGATT 762

Qy      79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp 98
           |||:::::::::         :::||||||   |||:::|||   :::||||||
Db     763 AGAGAGCGTTACTTGGCTGCAGTGCTTAGACAGGAGATTGCCTACTTTGATGATCTGGGC 822

Qy      99 IleGlyGluLeuAsnThrArgMet---ThrAspIleAspLysIleSerAspGlyIleGly 117
              ||||||:::   ||||||:::   ||||||         :::   :::|||
Db     823 GCCGGAGAAGTCGCCACTCGCATTCAGACCGATTGTCACCTTGTCCAAGAGGGAACATCT 882

Qy     118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
           :::|||:::||||||:::||||||               ||||||   |||      :::
Db     883 GAAAAGGTCGCTCTCGTATTCCAATATGCTGGTACTTTTGTCTGCGGTTTTGTTCTAGCT 942

Qy     138 LeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
              |||:::      :::|||               :::|||   |||:::||||||
Db     943 TTTGTCCGGTCACCTCGTCTTGCGGGCGCTCTTGTTTCCATCCTGCCAGTCATCATGCTC 1002

Qy     158 SerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr 177
                                 :::            :::      |||
Db    1003 TGCGGCGGTATTATGATGACGGCCATGGCCAAATATGGGACTGCGGCTCTTGATCACATC 1062

Qy     178 SerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197
           :::||||||||||::::::|||||||||||||:::   ||||||||||||||   ||||||
Db    1063 GCGAAAGCGGGCAGCTTGGCAGAAGAGGTCATAGGAAGTATCAGGACTGTTCAGGCTTTT 1122

Qy     198 ArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyrAlaTrp 217
              :::||||||   |||      |||
Db    1123 ---GGCAAGGAAAAGATCTTGGGTGACAAATTT--------------------------- 1152

Qy     218 PheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThrGlnAsn 237
                                                                     :::
Db    1153 ---------------------------------------------------GCCGATCAC 1161

Qy     238 LeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGly 257
           ::::::   :::|||      |||   |||   :::|||            |||
Db    1162 ATTGAGCAGAGCAAGATTGTCGGTAGAAAAGGCTCCATCTTTGAAGGCTTTGGTTTGAGC 1221

Qy     258 AlaValTyrPhePheMetAsnGlyThrThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeu 277
              ::::::|||   :::         |||   |||||||||:::||||||      |||
Db    1222 ATCATGTTCTTCGTCATCTACGCCGCTTATGCCCTCGCCTTCTTCTACGGTGGCATTCTC 1281
```

```
Qy     278 IleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerVal 297
            :::   ||||||:::            |||   |||:::  ||||||   |||:::
Db    1282 GTCAGTAACGGCCAAGCC------GACTCCGGTATCGTCATCAACGTTTTCATGTCCATC 1335

Qy     298 IleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAla 317
            :::         |||:::   :::            |||               |||
Db    1336 CTCATTGGTTCCTTCTCCATGGCGATGCTTGCTCCCGAGTTGGCTGCTGTCACCAAAGCG 1395

Qy     318 ArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPhe 337
           ||||||||||||         :::|||      ||||||:::   |||:::||||||:::
Db    1396 CGAGGCGCTGCTGCCAAGCTTTTCGCAACTATCGACCGTGTACCCGCCATTGATTCCGCC 1455

Qy     338 SerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSer 357
           |||      |||:::||||||:::   :::   |||   :::   |||:::||||||
Db    1456 AGCGAGGAAGGTTTCAAGCCCGATGGCCTTCGCGGTGAGATTAGCTTTGAAAATGTCAAG 1515

Qy     358 PheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLys 377
           |||:::||||||||||||||||||||||||   ||||||||||||               :::
Db    1516 TTCCATTATCCGTCCCGACCCAGCATTCCTATCTTGAAGGGCTTCACTACTACGTTTGAA 1575

Qy     378 SerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGln 397
           :::|||:::|||   ||||||||||||   :::||||||||||||||||||||||||
Db    1576 GCCGGTAAGACTTTCGCTCTCGTCGGAGCTAGTGGAAGCGGAAAGAGTACTGTCGTTTCT 1635

Qy     398 LeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIle 417
           |||::::::|||   |||||||||      |||   :::   :::|||      ||||||
Db    1636 CTTATTGAGAGGTTTTACGACCCCGTTAGTGGTGTTGTGAAGCTTGACGGTAGAGATATC 1695

Qy     418 ArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProVal 437
           |||:::||||||:::            |||      ||||||:::||||||||||||||
Db    1696 AGGTCGCTTAACCTCAATTGGCTTCGTCAACAGATTGGTCTTGTATCTCAGGAACCCACT 1755

Qy     438 LeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly---------------ArgAsp 452
           |||||||||||||||:::         |||:::::::::|||                  :::
Db    1756 CTTTTCGGTACCACCGTCCGCGGTAACGTCGAACACGGTCTCATCGGCTCCAGATATGAG 1815

Qy     453 AspValThrAspGluGlu------------MetGluArgAlaAlaArgGluAlaAsnAla 468
            :::   :::   ||||||            :::::::::|||      :::|||||||||
Db    1816 AATGCCTCTCTCGAGGAAAAATTTGAGCTTGTGAAGAAGGCTTGTGTCGATGCCAACGCC 1875

Qy     469 TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAla 488
           ::::::|||||||||:::   |||      ::::::|||:::|||||||||:::|||
Db    1876 CATAATTTTATCATGAAACTTCCGCAAGGTTATGATACCATGGTTGGTGAGCGCGGTATG 1935

Qy     489 GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro 508
              :::||||||||||||||||||||||:::||||||||||||||||:::|||      :::|||
Db    1936 CTGTTATCTGGTGGTCAGAAGCAGCGAGTTGCGATTGCTCGTGCTATCGTTTCCGACCCT 1995

Qy     509 LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 528
           :::||||||:::||||||||||||||||||||||||||||:::::::|||:::         |||
Db    1996 AGGATTTTGTTGTTGGACGAAGCTACCTCTGCTCTTGACACTCAGAGCGAAGGTATTGTG 2055

Qy     529 GlnAlaAlaLeuGluLys 534
           |||   ||||||:::|||
Db    2056 CAAGATGCTTTGGACAAG 2073
```

RESULT 13
US-08-612-734B-1
; Sequence 1, Application US/08612734B
; Patent No. 5914246
; GENERAL INFORMATION:
; APPLICANT: Peery, Robert B.
; APPLICANT: Skatrud, Paul L.
; APPLICANT: Tobin, Matthew B.
; TITLE OF INVENTION: Multiple Drug Resistance Gene of
; TITLE OF INVENTION: Aspergillus Fumigatus
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Eli Lilly and Company
; STREET: Lilly Corporate Center, DC1501
; CITY: Indianapolis
; STATE: Indiana
; COUNTRY: U.S.A.
; ZIP: 46285
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/612,734B
; FILING DATE: 08-MAR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Craig, Anne I.
; REGISTRATION NUMBER: 32,976
; REFERENCE/DOCKET NUMBER: X-9681
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-354-9570
; TELEFAX: 617-354-4043
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4047 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: other nucleic acid
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..4047
US-08-612-734B-1

```
Alignment Scores:
Pred. No.:                1.07e-99      Length:         4047
Score:                    935.50        Matches:        203
Percent Similarity:       57.01%        Conservative:   110
Best Local Similarity:    36.98%        Mismatches:     191
Query Match:              34.15%        Indels:         45
DB:                       2             Gaps:           7

US-09-873-409-7 (1-541) x US-08-612-734B-1 (1-4047)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            :::::::::   :::   ::::::         ||||||   |||||||||      :::
Db     331 ATACTCGTGAGCGCAATCTGTGCTATTGCTGCTGGTGCCGCTCTTCCCTTGTTCACAATT 390

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
            :::   |||   ::::::   :::         |||      :::   |||      |||
Db     391 CTCTTCGGTTCACTAGCC---TCAGCGTTCCAGGGCATATCTCTGGGTACTATGCCCTAT 447

Qy      41 SerPhePhe------ArgLeuThr---------LeuTyrTyrValGlyIleGlyValAla 55
                 |||      :::||||||         ||||||:::|||   :::|||:::|||
Db     448 CATGAGTTCTATCACAAACTGACTAAGAATGTGCTTTACTTTGTGTATCTCGGTATTGCC 507

Qy      56 AlaLeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThr 75
                 :::      |||:::               :::|||   |||            |||
Db     508 GAGTTTGTCACAGTCTATGTCAGCACCGTGGGTTTCATTTATACTGGCGAACATCTCACA 567

Qy      76 LysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAsp 95
           ::::::||||||:::   :::      ::::::|||   |||::::::      :::||||||
Db     568 CAGAAGATCCGTGAAAATTATCTTGAGGCTATCCTGAGGCAGAATATGGCTTACTTCGAC 627

Qy      96 SerCysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAsp 114
                    ||||||:::   ||||||:::|||   |||   :::   |||   |||
Db     628 AAGTTGGGCGCCGGTGAAGTTACCACGCGTATCACTGCTGATACCAACCTGATCCAGGAC 687

Qy     115 GlyIleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeu 134
              |||   :::|||:::   |||                  :::||||||
Db     688 GCCATCTCTGAGAAAGTTGGTCTCACTTTGACCGCATTCGCCACATTTGTAACCGCATTT 747

Qy     135 AlaValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeu 154
              |||      ||||||   |||||||||   |||:::      ||||||
Db     748 ATTGTCGCCTACGTCAAGTATTGGAAGTTGGCTCTGATCTGTACCTCAACTATCGTCGCG 807

Qy     155 IleMetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeu 174
           ::::::                  ||||||   ::::::      :::   |||   :::
```

```
Db    808 CTGGTCATGGTTATGGGAGGTGGGTCGAGGTTTATTGTGAAGTACAGCAAGAAATCTATT 867

Qy    175 SerAlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrVal 194
             :::|||         |||   ||||||||||||||:::|||||||||||
Db    868 GAAAGTTATGGTGCTGGTGGAACTGTCGCGGAAGAAGTCATCAGCTCCATTCGGAATGCT 927

Qy    195 IleAlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArg 214
             ||||||     |||:::|||      :::
Db    928 ACCGCTTTCGGCACTCAGGATAAGCTCGCCAAG--------------------------- 960

Qy    215 TyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyr 234
                                                                :::|||
Db    961 ---------------------------------------------------------CAATAC 966

Qy    235 ThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysVal 254
                :::|||    :::|||:::    :::|||:::|||:::    :::          :::
Db    967 GAAACCCATCTGGCTGAGGCTGAAAAATGGGGCGTCAAACAACAGGTCATCCTTGGTATG 1026

Qy    255 SerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGly 274
             :::|||   ::::::        |||          ||||||||   ||||||   |||
Db   1027 ATGATTGGTGGTATGTTCGGTATCATGTTCTCGAACTATGGTCTCGGTTTCTGGATGGGA 1086

Qy    275 ThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhe 294
          :::      ::::::      |||          :::|||   ||||||   |||
Db   1087 TCTCGCTTCGTTGTCGGTAAAGAA------GTCAACGTGGGCCAAGTTCTGACAGTTTTG 1140

Qy    295 PheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPhe 314
             |||::::::         |||:::   :::|||         |||:::   :::   |||
Db   1141 ATGTCTATCCTGATCGGTTCGTTCAGTTTGGGCAACGTCGCCCCCAATGGTCAGGCCTTT 1200

Qy    315 AlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIle 334
                      ||||||      |||:::         ||||||::::::          :::
Db   1201 ACGAATGGTGTTGCTGCGGCCGCGAAGATTTACAGCACGATTGACCGCAGATCGCCACTG 1260

Qy    335 AspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLys 354
          |||   :::|||      |||            :::      ||||||   :::||||||:::
Db   1261 GACCCCTATTCTGACGAAGGGAAGGTACTCGACCATTTTGAAGGAAATATCGAATTTCGC 1320

Qy    355 AsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeu 374
          ||||||         |||||||||||||||   :::   :::::::::   ::::::|||
Db   1321 AATGTCAAACACATCTACCCTTCAAGACCCGAAGTTACAGTCATGGAAGATGTCTCTTTA 1380

Qy    375 ArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThr 394
             :::   :::|||:::|||   ||||||||||||   :::|||||||||||||||||||
Db   1381 TCGATGCCCGCCGGAAAGACTACCGCATTGGTGGGCCCATCTGGCTCTGGAAAGAGTACT 1440

Qy    395 ValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGlu 414
          ||||||   |||::::::|||   |||   |||      |||   :::::::::|||
Db   1441 GTTGTCGGCTTGGTGGAGCGCTTTTACCTTCCAGTAGGAGGCCAGGTATTGCTGGACGGC 1500

Qy    415 AsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGln 434
          :::||||||:::   ||||||:::|||      |||      |||   :::||||||||||
Db   1501 CATGATATCCAAACCCTCAACCTCCGTTGGCTGCGACAGCAGATCTCTCTTGTCAGCCAG 1560

Qy    435 GluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly------------ 450
          |||||||||||||||   ||||||||||   ||||||::::::|||
Db   1561 GAACCTGTTCTTTTCAGCACCACGATCTTTAGAAACATCGAACATGGCTTGATTGGCACC 1620

Qy    451 ---------------ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 465
                         |||         |||   :::|||   |||||||||
Db   1621 AAATTCGAGCATGAGTCGAAGGACAAGATTAGAGAGCTCGTTGAGAATGCGGCCAGAATG 1680

Qy    466 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 485
          |||||||||:::||||||||||||          |||      ::::::|||   ||||||:::
Db   1681 GCCAATGCTCATGATTTTATTATGGCTCTGCCTGAAGGTTACGATACGAATGTGGGTCAG 1740

Qy    486 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 505
          :::|||      :::||||||||||||||||||||||||||||||||||||:::|||
Db   1741 CGTGGTTTCTTACTTTCAGGAGGTCAGAAGCAACGTATTGCCATTGCTCGTGCCATTGTC 1800

Qy    506 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 525
             :::||||||||||:::|||||||||||||||||||||||||::::::|||:::
Db   1801 AGTGACCCCAAGATTCTGTTGCTTGATGAAGCTACATCAGCTTTGGATACCAAGTCCGAG 1860

Qy    526 SerAlaValGlnAlaAlaLeuGluLys 534
                ||||||||||||||:::|||
Db   1861 GGCGTCGTCCAAGCCGCTCTTGATAAA 1887
```

RESULT 14
US-08-612-734B-3
; Sequence 3, Application US/08612734B
; Patent No. 5914246
; GENERAL INFORMATION:
;   APPLICANT: Peery, Robert B.
;   APPLICANT: Skatrud, Paul L.
;   APPLICANT: Tobin, Matthew B.
;   TITLE OF INVENTION: Multiple Drug Resistance Gene of
;   TITLE OF INVENTION: Aspergillus Fumigatus
;   NUMBER OF SEQUENCES: 9
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Eli Lilly and Company
;   STREET: Lilly Corporate Center, DC1501
;   CITY: Indianapolis
;   STATE: Indiana
;   COUNTRY: U.S.A.
;   ZIP: 46285
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/612,734B
;   FILING DATE: 08-MAR-1996
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Craig, Anne I.
;   REGISTRATION NUMBER: 32,976
;   REFERENCE/DOCKET NUMBER: X-9681
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 617-354-9570
;   TELEFAX: 617-354-4043
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4800 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: other nucleic acid
US-08-612-734B-3

```
Alignment Scores:
Pred. No.:              2.86e-97      Length:         4800
Score:                  916.00        Matches:        202
Percent Similarity:     55.47%        Conservative:   102
Best Local Similarity:  36.86%        Mismatches:     182
Query Match:            33.44%        Indels:         62
DB:                     2             Gaps:           7

US-09-873-409-7 (1-541) x US-08-612-734B-3 (1-4800)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          ::::::   |||   |||||||||         |||         |||      ||||||
Db    693 ATTCTCTTCGGTTCACTAGCCTCAGCGTTCCAGGGCATATCTCTGGGTACTATGCCCTAT 752

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
                                                                   :::
Db    753 ---------------------------------------------------------CAT 755

Qy     41 SerPhePhe---ArgLeuThr---------LeuTyrTyrValGlyIleGlyValAlaAla 56
             |||:::   :::||||||         ||||||:::|||   :::|||:::|||
Db    756 GAGTTCTATCACAAACTGACTAAGAATGTGCTTTACTTGTGTATCTCGGTATTGCCGAG 815

Qy     57 LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys 76
             :::      |||:::            :::|||   |||            |||:::
Db    816 TTTGTCACAGTCTATGTCAGCACCGTGGGTTTCATTTATACTGGCGAACATCTCACACAG 875
```

```
Qy       77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96
            :::||||||:::   :::   ::::::|||  |||::::::   :::||||||
Db      876 AAGATCCGTGAAAATTATCTTGAGGCTATCCTGAGGCAGAATATGGCTTACTTCGACAAG 935

Qy       97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
                     ||||||:::   ||||||:::|||   |||  :::   |||   |||
Db      936 TTGGGCGCCGGTGAAGTTACCACGCGTATCACTGCTGATACCAACCTGATCCAGGACGCC 995

Qy      116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
            |||   :::|||:::   |||                     :::||||||
Db      996 ATCTCTGAGAAAGTTGGTCTCACTTTGACCGCATTCGCCACATTTGTAACCGCATTTATT 1055

Qy      136 ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
            |||      ||||||   |||||||||   |||:::      ||||||         :::
Db     1056 GTCGCCTACGTCAAGTATTGGAAGTTGGCTCTGATCTGTACCTCAACTATCGTCGCGCTG 1115

Qy      156 MetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
            :::                  ||||||   ::::::      :::   |||   :::
Db     1116 GTCATGGTTATGGGAGGTGGGTCGAGGTTTATTGTGAAGTACAGCAAGAAATCTATTGAA 1175

Qy      176 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIle 195
            :::|||         |||   |||||||||||||||:::||||||||||||
Db     1176 AGTTATGGTGCTGGTGGAACTGTCGCGGAAGAAGTCATCAGCTCCATTCGGAATGCTACC 1235

Qy      196 AlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeuLeuAsnIleThrArgTyr 215
            ||||||      |||:::|||      :::
Db     1236 GCTTTCGGCACTCAGGATAAGCTCGCCAAG------------------------------ 1265

Qy      216 AlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu***PheValArgTyrThr 235
                                                            :::|||
Db     1266 ---------------------------------------------------CAATACGAA 1274

Qy      236 GlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSer 255
               :::|||   :::|||:::   :::|||:::|||:::            :::
Db     1275 ACCCATCTGGCTGAGGCTGAAAAATGGGGCGTCAAACAACAGGTCATCCTTGGTATGATG 1334

Qy      256 LeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThr 275
            :::|||   ::::::      |||         |||||||||   ||||||   |||:::
Db     1335 ATTGGTGGTATGTTCGGTATCATGTTCTCGAACTATGGTCTCGGTTTCTGGATGGGATCT 1394

Qy      276 SerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePhe 295
                  ::::::   |||         :::|||   ||||||   |||
Db     1395 CGCTTCGTTGTCGGTAAAGAA------GTCAACGTGGGCCAAGTTCTGACAGTTTTGATG 1448

Qy      296 SerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAla 315
            |||::::::      |||:::   :::|||         |||:::   :::   |||
Db     1449 TCTATCCTGATCGGTTCGTTCAGTTTGGGCAACGTCGCCCCCAATGGTCAGGCCTTTACG 1508

Qy      316 IleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAsp 335
                              ||||||   |||:::   ||||||::::::      :::|||
Db     1509 AATGGTGTTGCTGCGGCCGCGAAGATTTACAGCACGATTGACCGCAGATCGCCACTGGAC 1568

Qy      336 AsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsn 355
               :::|||   |||      :::      ||||||   :::||||||:::|||
Db     1569 CCCTATTCTGACGAAGGGAAGGTACTCGACCATTTTGAAGGAAATATCGAATTTCGCAAT 1628

Qy      356 ValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArg 375
            |||         ||||||||||||||   :::   :::::::::   ::::::|||
Db     1629 GTCAAACACATCTACCCTTCAAGACCCGAAGTTACAGTCATGGAAGATGTCTCTTTATCG 1688

Qy      376 IleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrVal 395
            :::   :::|||:::|||   ||||||||||||   :::||||||||||||||||||||||
Db     1689 ATGCCCGCCGGAAAGACTACCGCATTGGTGGGCCCATCTGGCTCTGGAAAGAGTACTGTT 1748

Qy      396 ValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsn 415
            |||   |||::::::|||   |||   |||      |||   :::::::::|||   :::
Db     1749 GTCGGCTTGGTGGAGCGCTTTTACCTTCCAGTAGGAGGCCAGGTATTGCTGGACGGCCAT 1808

Qy      416 AspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGlu 435
            ||||||:::   ||||||:::|||      |||      |||   :::||||||||||||
Db     1809 GATATCCAAACCCTCAACCTCCGTTGGCTGCGACAGCAGATCTCTCTTGTCAGCCAGGAA 1868

Qy      436 ProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly--------------- 450
            ||||||||||||   |||||||||         ||||||::::::|||
Db     1869 CCTGTTCTTTTCAGCACCACGATCTTTAGAAACATCGAACATGGCTTGATTGGCACCAAA 1928

Qy      451 ------------ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAla 466
                        |||         |||   :::|||   |||||||||   |||
Db     1929 TTCGAGCATGAGTCGAAGGACAAGATTAGAGAGCTCGTTGAGAATGCGGCCAGAATGGCC 1988

Qy      467 AsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLys 486
            ||||||:::||||||||||||         |||         ::::::|||   ||||||::::::
Db     1989 AATGCTCATGATTTTATTATGGCTCTGCCTGAAGGTTACGATACGAATGTGGGTCAGCGT 2048

Qy      487 GlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArg 506
            |||         :::|||||||||||||||||||||||||||||||||||||||:::|||
Db     2049 GGTTTCTTACTTTCAGGAGGTCAGAAGCAACGTATTGCCATTGCTCGTGCCATTGTCAGT 2108

Qy      507 AsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSer 526
            :::||||||||||||:::||||||||||||||||||||||||::::::|||:::
Db     2109 GACCCCAAGATTCTGTTGCTTGATGAAGCTACATCAGCTTTGGATACCAAGTCCGAGGGC 2168

Qy      527 AlaValGlnAlaAlaLeuGluLys 534
                  |||||||||||||||:::|||
Db     2169 GTCGTCCAAGCCGCTCTTGATAAA 2192
```

RESULT 15

```
US-08-996-545-1
; Sequence 1, Application US/08996545
; Patent No. 5928898
;  GENERAL INFORMATION:
;    APPLICANT:  Skatrud, Paul L.
;    APPLICANT:  de Waard, Maarten A.
;    APPLICANT:  Peery, Robert B.
;    APPLICANT:  Andrade, Alan C.
;    TITLE OF INVENTION:  Multiple Drug Resistance Gene atrD of
;    TITLE OF INVENTION:  Aspergillus Nidulans
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  U.S.
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/996,545
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Webster, Thomas D.
;      REGISTRATION NUMBER:  39,872
;      REFERENCE/DOCKET NUMBER:  X-11766
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-3334
;      TELEFAX:  317-276-2763
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4002 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..4002
US-08-996-545-1
```

```
Alignment Scores:
Pred. No.:              3.62e-97      Length:         4002
Score:                  914.00        Matches:        201
Percent Similarity:     56.93%        Conservative:   103
Best Local Similarity:  37.64%        Mismatches:     181
Query Match:            33.37%        Indels:         50
DB:                     2             Gaps:           6

US-09-873-409-7 (1-541) x US-08-996-545-1 (1-4002)

Qy       15 LeuProLeuMetProLeuValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu 34
               :::||||||:::|||       |||                           |||
Db      368 GTGCCATTGCTGCCGCGTCGACTTTCCAGA------------------GGATAATGTTAT 409

Qy       35 ValGlnThrAsnThrTyrSerPhePhe------------ArgLeuThrLeuTyrTyrVal 50
            :::::::::    |||     |||:::                 :::      ||||||:::|||
Db      410 ATCAAATCTCGTACGAC-GAGTTCTATGATGAATTGACCAAGAACGTACTGTACTTCGTA 468

Qy       51 GlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThr 70
               :::|||:::        :::        |||:::                :::|||   |||
Db      469 TACCTCGGTATCGGCGAGTTTGTCACTGTCTATGTTAGTACTGTTGGCTTCATCTATACC 528

Qy       71 AlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAsp 90
                           |||::::::||||||:::     :::        |||:::|||   |||:::
Db      529 GGAGAACACGCCACGCAGAAGATCCGCGAGTATTACCTTGAGTCTATCCTGCGCCAGAAC 588

Qy       91 IleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMetThr---AspIle 109
            ||||||:::||||||             |||||||:::     ||||||:::|||   |||
Db      589 ATTGGCTATTTTGATAAACTCGGTGCCGGGGAAGTGACCACCCGTATAACAGCCGATACA 648

Qy      110 AspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThr 129
            :::     |||     ||||||||   :::|||:::     |||        ::::::|||
Db      649 AACCTTATCCAGGATGGCATTTCGGAGAAGGTCGGTCTCACTTTGACTGCCCTGGCGACA 708

Qy      130 PheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeu 149
            |||                 :::        ||||||   |||||||||   |||:::
Db      709 TTCGTGACAGCATTCATTATCGCCTACGTCAAATACTGGAAGTTGGCTCTAATTTGCAGC 768

Qy      150 SerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeu 169
            ||||||          ::::::   :::                |||:::   :::|||
Db      769 TCAACAATTGTGGCCCTCGTTCTCACCATGGCGGTGGTTCTCAGTTTATCATCAAGTAC 828

Qy      170 ThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSer 189
            :::     |||   |||   :::|||        |||   ||||||||||||||:::|||
Db      829 AGCAAAAAGTCGCTTGACAGCTACGGTGCAGGCGGCACTGTTGCGGAAGAGGTCATCAGC 888

Qy      190 SerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgSerPheLeu 209
            |||||||||         ||||||   |||:::|||      :::
Db      889 TCCATCAGAAATGCCACAGCGTTTGGCACCCAAGACAAGCTTGCGAAG------------ 936

Qy      210 LeuAsnIleThrArgTyrAlaTrpPheTyrPheProGlnTrpLeuLeuSerCysValLeu 229

Db      936 ------------------------------------------------------------ 936

Qy      230 ***PheValArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThr 249
                      :::|||   :::|||   :::|||:::   :::|||   |||
Db      937 ---------CAGTATGAGGTCCACTTAGACGAAGCTGAGAAATGGGGAACAAAGAACCAG 987

Qy      250 IleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeu 269
            |||                 :::||||||::::::        |||           |||||||||
Db      988 ATTGTCATGGGTTTCATGATTGGCGCCATGTTTGGCCTTATGTACTCGAACTACGGTCTT 1047

Qy      270 AlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThr 289
               ||||||   |||:::      :::::::::|||          :::|||
Db     1048 GGCTTCTGGATGGGTTCTCGTTTCCTGGTAGATGGT------GCAGTCGATGTGGGTGAT 1101

Qy      290 ValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValPro 309
            :::|||   |||      :::::::::        |||:::   :::|||        |||
Db     1102 ATTCTCACAGTTCTCATGGCCATCTTGATCGGATCGTTCTCCTTGGGGAACGTTAGTCCA 1161

Qy      310 HisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAsp 329
            :::   :::   |||      |||        ||||||      ||||||      ||||||
Db     1162 AATGCTCAAGCATTTACAAACGCTGTGGCCGCGGCCGCAAAGATATTTGGAACGATCGAT 1221

Qy      330 LysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGly 349
            ::::::      :::|||   :::|||      |||          :::      ||||||
Db     1222 CGCCAGTCCCCATTAGATCCATATTCGAACGAAGGGAAGACGCTCGACCATTTTGAGGGC 1281

Qy      350 ThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeu 369
               :::|||   :::||||||               ||||||||||||||   :::   ::::::
Db     1282 CACATTGAGTTACGCAATGTCAAGCATATTTACCCATCTAGACCCGAGGTCACCGTCATG 1341

Qy      370 LysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGly 389
            :::     ::::::|||   :::   :::|||:::|||   ||||||||||||   :::|||
Db     1342 GAGGATGTTTCTCTGTCAATGCCCGCTGGAAAAACAACCGCTTTAGTCGGCCCCTCTGGC 1401

Qy      390 SerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPhe 409
            ||||||||||||||||||||||||   |||::::::|||   |||   |||      |||
Db     1402 TCTGGAAAAAGTACGGTGGTCGGCTTGGTTGAGCGATTCTACATGCCTGTTCGCGGTACG 1461

Qy      410 IleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle 429
            :::::::::|||   :::|||||||:::   ||||||:::|||         |||      |||
Db     1462 GTTTTGCTGGATGGCCATGACATCAAGGACCTCAATCTCCGCTGGCTTCGCCAACAGATC 1521

Qy      430 GlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr 449
               :::||||||||||||||||||||||||||||||||||||   ||||||::::::
Db     1522 TCTTTGGTTAGCCAGGAGCCTGTTCTTTTTGGCACGACGATTTATAAGAATATTAGGCAC 1581

Qy      450 Gly---------------------------ArgAspAspValThrAspGluGluMetGlu 460
            |||                              :::|||      |||   :::|||
Db     1582 GGTCTCATCGGCACAAAGTACGAGAATGAATCCGAGGATAAGGTCCGGGAACTCATCGAG 1641

Qy      461 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 480
               ||||||:::   |||||||||:::||||||||         |||      :::
Db     1642 AACGCGGCAAAAATGGCGAATGCTCATGACTTTATTACTGCCTTGCCTGAAGGTTATGAG 1701

Qy      481 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 500
            |||   ||||||::::::|||      :::||||||||||||||||||||||||||||||
Db     1702 ACCAATGTTGGGCAGCGTGGCTTTCTCCTTTCAGGTGGCCAGAAACAGCGCATTGCAATC 1761

Qy      501 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 520
            |||||||||:::|||   :::|||||||||||:::|||||||||||||||||||||||||
Db     1762 GCCCGTGCCGTTGTTAGTGACCCAAAAATCCTGCTCCTGGATGAAGCTACTTCGGCCTTG 1821

Qy      521 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 534
            |||::::::|||:::         |||||||||||||||:::
Db     1822 GACACAAAATCCGAAGGCGTGGTTCAAGCAGCTTTGGAGAGG 1863

Search completed: March 31, 2003, 15:37:48
Job time : 148.227 secs
```

```
                 GenCore version 5.1.4_p5_4578
          Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 31, 2003, 13:52:37 ; Search time 187.092 Seconds
                                           (without alignments)
                                          2337.888 Million cell updates/sec

Title:          US-09-873-409-8
Perfect score:  2592
Sequence:       1 MILGILASLVNGACLPLMPL..........SESKSAVQAALEKDTPRYSF 514

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       574371 seqs, 425486471 residues

Total number of hits satisfying chosen parameters:      1148742

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115424_19290/app_query.fasta_1.7544
-DB=Published_Applications_NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09873409_@CGN_1_1_960_@runat_27032003_115424_19290
-NCPU=6 -ICPU=3 -NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Published_Applications_NA:*
                1:  /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
                2:  /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
                3:  /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
                4:  /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
                5:  /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
                6:  /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
                7:  /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
                8:  /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
                9:  /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
                10:  /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
                11:  /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
                12:  /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
                13:  /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
                14:  /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

      Pred. No. is the number of results predicted by chance to have a
      score greater than or equal to the score of the result being printed,
      and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2592	100.0	1940	10	US-09-873-409-16	Sequence 16, Appl
2	2568.5	99.1	2021	10	US-09-873-409-15	Sequence 15, Appl
3	2552	98.5	3621	10	US-09-873-409-14	Sequence 14, Appl
4	2528.5	97.6	3702	10	US-09-873-409-13	Sequence 13, Appl
5	1699.5	65.6	3177	10	US-09-873-409-12	Sequence 12, Appl
6	1590	61.3	4189	10	US-09-866-866A-5	Sequence 5, Appli
7	1579	60.9	3860	10	US-09-866-866A-1	Sequence 1, Appli
8	1576	60.8	4317	9	US-10-044-671-1	Sequence 1, Appli
9	1571.5	60.6	4254	10	US-09-917-800A-1424	Sequence 1424, Ap
10	1570	60.6	3860	10	US-09-866-866A-3	Sequence 3, Appli
11	1570	60.6	8630	10	US-09-306-417-1	Sequence 1, Appli
12	1570	60.6	8630	10	US-09-306-417-2	Sequence 2, Appli
13	1566.5	60.4	4425	10	US-09-769-097-3	Sequence 3, Appli
14	1564.5	60.4	4369	10	US-09-769-097-1	Sequence 1, Appli
15	1562.5	60.3	4643	9	US-10-072-621-2	Sequence 2, Appli
16	1559.5	60.2	4788	10	US-09-866-866A-7	Sequence 7, Appli
17	1479	57.1	3924	10	US-09-880-107-2299	Sequence 2299, Ap
18	1445	55.7	3912	10	US-09-917-800A-1560	Sequence 1560, Ap
19	1243.5	48.0	5010	10	US-09-917-800A-483	Sequence 483, App
20	1141.5	44.0	1175	10	US-09-873-409-11	Sequence 11, Appl
21	1101.5	42.5	2856	10	US-09-873-409-10	Sequence 10, Appl
22	1050.5	40.5	1810	10	US-09-749-340-4	Sequence 4, Appli
23	1050.5	40.5	4175	10	US-09-749-340-3	Sequence 3, Appli
24	1017	39.2	3861	9	US-09-938-842A-2237	Sequence 2237, Ap
25	965.5	37.2	3861	9	US-09-938-842A-263	Sequence 263, App
26	937.5	36.2	4002	9	US-09-758-828-1	Sequence 1, Appli
27	937.5	36.2	4002	9	US-09-758-828-3	Sequence 3, Appli
28	915	35.3	4653	9	US-10-101-388-2	Sequence 2, Appli
29	828	31.9	2066	10	US-09-873-409-9	Sequence 9, Appli
30	809	31.2	3792	9	US-09-882-694-10	Sequence 10, Appl
31	766.5	29.6	8036	9	US-10-101-388-1	Sequence 1, Appli
32	755.5	29.1	2681	10	US-09-749-340-8	Sequence 8, Appli
33	744.5	28.7	3512	10	US-09-749-340-7	Sequence 7, Appli
34	739.5	28.5	2698	10	US-09-749-340-5	Sequence 5, Appli
35	712	27.5	3999	9	US-09-882-694-9	Sequence 9, Appli
36	668	25.8	2298	9	US-10-156-239-18	Sequence 18, Appl
37	668	25.8	2298	10	US-09-795-693-18	Sequence 18, Appl
38	668	25.8	3408	9	US-10-156-239-16	Sequence 16, Appl
39	668	25.8	3408	10	US-09-795-693-16	Sequence 16, Appl
40	668	25.8	3512	9	US-10-072-621-1	Sequence 1, Appli
41	619.5	23.9	2674	10	US-09-917-800A-1500	Sequence 1500, Ap
42	607	23.4	1749	10	US-09-815-242-9908	Sequence 9908, Ap
43	603	23.3	2247	9	US-09-870-759-40	Sequence 40, Appl
44	602	23.2	1749	10	US-09-815-242-6018	Sequence 6018, Ap
45	598	23.1	2824	9	US-09-954-531-170	Sequence 170, App

ALIGNMENTS

```
RESULT 1
US-09-873-409-16
; Sequence 16, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
;   LENGTH: 1940
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-409-16

Alignment Scores:
Pred. No.:              1.14e-267      Length:        1940
Score:                  2592.00        Matches:       514
Percent Similarity:     100.00%        Conservative:  0
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            100.00%        Indels:        0
DB:                     10             Gaps:          0
```

```
US-09-873-409-8 (1-514) x US-09-873-409-16 (1-1940)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      34 ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 93

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153

Qy      41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     154 TCTTTCTTCAGGTTGACCCTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGT 213

Qy      61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     214 TACATACAGATTTCCTTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAA 273

Qy      81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     274 CAGTTTTTTCATTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333

Qy     101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     334 GAACTTAACACTCGCATGACAGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATT 393

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAG 453

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     454 GGCTGGAAACTCACCCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA 513

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAG 633

Qy     201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     634 GAGAAAGAACTTCAAAGGTATACACAGAATCTCAAAGATGCAAAGGATTTTGGCATAAAA 693

Qy     221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     694 AGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAATGGAACCTAT 753

Qy     241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     754 GGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGATATACCATC 813

Qy     261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     814 GGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATTGGAGCAGCA 873

Qy     281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     874 GTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTT 933

Qy     301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     934 ATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATA 993

Qy     321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     994 GAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAG 1053

Qy     341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1054 ATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTC 1113

Qy     361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1114 AATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGAT 1173

Qy     381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1174 GGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGAC 1233

Qy     401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1234 CATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATC 1293

Qy     421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1294 AAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAAT 1353

Qy     441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1354 GCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGA 1413

Qy     461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1414 GCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAAC 1473

Qy     481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1474 CCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCT 1533

Qy     501 ValGlnAlaAlaLeuGluLysAspThrProArgTyrSerPhe 514
           ||||||||||||||||||||||||||||||||||||||||||
Db    1534 GTTCAAGCTGCACTGGAGAAGGATACCCCCAGGTATTCATTT 1575

RESULT 2
US-09-873-409-15
; Sequence 15, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION:  A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 15
;   LENGTH: 2021
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: No. US20020037522A1e
;   LOCATION: (723)..(723)
;   OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-15

Alignment Scores:
Pred. No.:               3.97e-265     Length:        2021
Score:                   2568.50       Matches:       514
Percent Similarity:      95.01%        Conservative:  0
Best Local Similarity:   95.01%        Mismatches:    0
Query Match:             99.09%        Indels:        27
DB:                      10            Gaps:          1

US-09-873-409-8 (1-514) x US-09-873-409-15 (1-2021)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      34 ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 93

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db      94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153

Qy      41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     154 TCTTTCTTCAGGTTGACCCTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGT 213

Qy      61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     214 TACATACAGATTTCCTTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAA 273

Qy      81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     274 CAGTTTTTTCATTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333

Qy     101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     334 GAACTTAACACTCGCATGACAGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATT 393

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAG 453

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     454 GGCTGGAAACTCACCCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA 513

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAG 633

Qy     201 GluLysGluLeuGln--------------------------------------------- 205
              |||||||||||||||
Db     634 GAGAAAGAACTTCAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTTGGTTTTATTTT 693

Qy     206 ------------------------------------ArgTyrThrGlnAsnLeuLysAsp 213
                                                  ||||||||||||||||||||||||
Db     694 CCCCAGTGGCTACTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGAATCTCAAAGAT 753

Qy     214 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 233
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     754 GCAAAGGATTTTGGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTAC 813

Qy     234 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 253
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     814 TTCTTTATGAATGGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAAT 873

Qy     254 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 273
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     874 GGAGAACCTGGATATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGC 933

Qy     274 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 293
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     934 AGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCT 993

Qy     294 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 313
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     994 GCCTTTCATATTTTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCT 1053

Qy     314 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 333
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1054 GGATATAAACCTGAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTAT 1113

Qy     334 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 353
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1114 CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAG 1173

Qy     354 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 373
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1174 ACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG 1233

Qy     374 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 393
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1234 AGGTTATATGATCCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1293

Qy     394 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 413
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1294 AATGTGCGGCATTATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGG 1353

Qy     414 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 433
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1354 ACCACCATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1413

Qy     434 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 453
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1414 AGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAAT 1473

Qy     454 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 473
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1474 ACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATT 1533

Qy     474 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 493
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1534 GCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTG 1593

Qy     494 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAspThrProArgTyrSer 513
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1594 GATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTGGAGAAGGATACCCCCAGGTATTCA 1653

Qy     514 Phe 514
              |||
Db    1654 TTT 1656
```

```
RESULT 3
US-09-873-409-14
; Sequence 14, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
;  APPLICANT: Frank, Markus
;  APPLICANT:  Sayegh, Mohamed
;  TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
;  TITLE OF INVENTION:  Homologue on Chromosome 7p15-21 and Uses Thereof
;  FILE REFERENCE: 81994/268611
;  CURRENT APPLICATION NUMBER: US/09/873,409
;  CURRENT FILING DATE:  2001-06-05
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14
;   LENGTH: 3621
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-873-409-14

Alignment Scores:
Pred. No.:              5.61e-263     Length:          3621
Score:                  2552.00       Matches:         507
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            98.46%        Indels:          0
DB:                     10            Gaps:            0

US-09-873-409-8 (1-514) x US-09-873-409-14 (1-3621)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      34 ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 93

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153

Qy      41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     154 TCTTTCTTCAGGTTGACCCTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGT 213

Qy      61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     214 TACATACAGATTTCCTTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAA 273

Qy      81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     274 CAGTTTTTCATTCAGTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333

Qy     101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     334 GAACTTAACACTCGCATGACAGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATT 393

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAG 453

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     454 GGCTGGAAACTCACCCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA 513

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAG 633

Qy     201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     634 GAGAAAGAACTTCAAAGGTATACACAGAATCTCAAAGATGCAAAGGATTTTGGCATAAAA 693

Qy     221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     694 AGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAATGGAACCTAT 753

Qy     241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     754 GGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGATATACCATC 813

Qy     261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     814 GGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATTGGAGCAGCA 873

Qy     281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     874 GTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTT 933

Qy     301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     934 ATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATA 993

Qy     321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     994 GAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAG 1053

Qy     341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1054 ATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTC 1113

Qy     361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1114 AATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGAT 1173

Qy     381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1174 GGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGAC 1233

Qy     401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1234 CATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATC 1293

Qy     421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1294 AAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAAT 1353

Qy     441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1354 GCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGA 1413

Qy     461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1414 GCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAAC 1473

Qy     481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1474 CCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCT 1533

Qy     501 ValGlnAlaAlaLeuGluLys 507
           |||||||||||||||||||||
Db    1534 GTTCAAGCTGCACTGGAGAAG 1554
```

RESULT 4
US-09-873-409-13
; Sequence 13, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13
;   LENGTH: 3702
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: No. US20020037522A1e
;   LOCATION: (723)..(723)
;   OTHER INFORMATION: n at position 723 represents any nucleotide (A, T, C or G)
US-09-873-409-13

```
Alignment Scores:
Pred. No.:              1.9e-260      Length:        3702
Score:                  2528.50       Matches:       507
Percent Similarity:     94.94%        Conservative:  0
Best Local Similarity:  94.94%        Mismatches:    0
Query Match:            97.55%        Indels:        27
DB:                     10            Gaps:          1

US-09-873-409-8 (1-514) x US-09-873-409-13 (1-3702)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      34 ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 93

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153

Qy      41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     154 TCTTTCTTCAGGTTGACCCTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGT 213

Qy      61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     214 TACATACAGATTTCCTTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAA 273

Qy      81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     274 CAGTTTTTCATTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333

Qy     101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     334 GAACTTAACACTCGCATGACAGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATT 393

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAG 453

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     454 GGCTGGAAACTCACCCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA 513

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAG 633

Qy     201 GluLysGluLeuGln--------------------------------------------- 205
           |||||||||||||||
Db     634 GAGAAAGAACTTCAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTTGGTTTTATTTT 693

Qy     206 -------------------------------------ArgTyrThrGlnAsnLeuLysAsp 213
                                                ||||||||||||||||||||||||
Db     694 CCCCAGTGGCTACTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGAATCTCAAAGAT 753

Qy     214 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 233
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     754 GCAAAGGATTTTGGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTAC 813

Qy     234 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 253
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     814 TTCTTTATGAATGGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAAT 873

Qy     254 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 273
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     874 GGAGAACCTGGATATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGC 933

Qy     274 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 293
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     934 AGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCT 993

Qy     294 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 313
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     994 GCCTTTCATATTTTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCT 1053

Qy     314 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 333
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1054 GGATATAAACCTGAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTAT 1113

Qy     334 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 353
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1114 CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAG 1173

Qy     354 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 373
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1174 ACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG 1233

Qy     374 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 393
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1234 AGGTTATATGATCCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1293

Qy     394 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 413
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1294 AATGTGCGGCATTATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGG 1353

Qy     414 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 433
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1354 ACCACCATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1413

Qy     434 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 453
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1414 AGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAAT 1473

Qy     454 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 473
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1474 ACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATT 1533

Qy     474 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 493
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1534 GCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTG 1593

Qy     494 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 507
           ||||||||||||||||||||||||||||||||||||||||||
Db    1594 GATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTGGAGAAG 1635
```

RESULT 5
US-09-873-409-12
; Sequence 12, Application US/09873409
; Patent No. US20020037522A1
; GENERAL INFORMATION:
; APPLICANT: Frank, Markus
; APPLICANT: Sayegh, Mohamed
; TITLE OF INVENTION: A Gene Encoding a Multidrug Resistance Human P-Glycoprotein
; TITLE OF INVENTION: Homologue on Chromosome 7p15-21 and Uses Thereof
; FILE REFERENCE: 81994/268611
; CURRENT APPLICATION NUMBER: US/09/873,409
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 12
;   LENGTH: 3177
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: No. US20020037522A1e
;   LOCATION: (198)..(198)
;   OTHER INFORMATION: n at position 198 represents any nucleotide (A, T, C or G)
US-09-873-409-12

```
Alignment Scores:
Pred. No.:              8.09e-172     Length:         3177
Score:                  1699.50       Matches:        343
Percent Similarity:     92.70%        Conservative:   0
Best Local Similarity:  92.70%        Mismatches:     0
Query Match:            65.57%        Indels:         27
DB:                     10            Gaps:           1

US-09-873-409-8 (1-514) x US-09-873-409-12 (1-3177)

Qy     165 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 184
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCA 60

Qy     185 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 204
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAGGAGAAAGAACTT 120

Qy     205 Gln--------------------------------------------------------- 205
           |||
Db     121 CAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTTGGTTTTATTTTCCCCAGTGGCTA 180

Qy     206 ------------------------ArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 217
                                   ||||||||||||||||||||||||||||||||||||
Db     181 CTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGAATCTCAAAGATGCAAAGGATTTT 240

Qy     218 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 237
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAAT 300

Qy     238 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 257
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     301 GGAACCTATGGACTTGCTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGA 360

Qy     258 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 277
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 TATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATT 420

Qy     278 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 297
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATT 480

Qy     298 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 317
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCT 540

Qy     318 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 337
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCA 600

Qy     338 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 357
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 TCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTG 660

Qy     358 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 377
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 GTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 720

Qy     378 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 397
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 CCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCAT 780

Qy     398 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 417
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 TATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGT 840

Qy     418 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 437
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 AACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGG 900

Qy     438 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 457
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     901 GAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGG 960

Qy     458 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 477
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     961 GAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTA 1020

Qy     478 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 497
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1021 GTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGC 1080

Qy     498 LysSerAlaValGlnAlaAlaLeuGluLys 507
           ||||||||||||||||||||||||||||||
Db    1081 AAGTCAGCTGTTCAAGCTGCACTGGAGAAG 1110
```

RESULT 6
US-09-866-866A-5
; Sequence 5, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
;   LENGTH: 4189
;   TYPE: DNA
;   ORGANISM: Mus musculus
US-09-866-866A-5

```
Alignment Scores:
Pred. No.:              6.47e-160     Length:         4189
Score:                  1590.00       Matches:        307
Percent Similarity:     76.32%        Conservative:   99
Best Local Similarity:  57.71%        Mismatches:     94
Query Match:            61.34%        Indels:         32
DB:                     10            Gaps:           5

US-09-873-409-8 (1-514) x US-09-866-866A-5 (1-4189)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||||||||||    ||||||::::::::::::|||         ||||||||:::    |||
Db     148 ATGATTCTGGGAACTCTCGCTGCTATTATCCATGGAACATTACTTCCCCTCTTGATGCTG 207

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
           |||    |||    |||:::|||:::
Db     208 GTGTTTGGAAACATGACAGATAGTTTTACAAAAGCAGAAGCCAGTATTCTGCCAAGCATT 267

Qy      31 ------------------SerGlyCysLeuValGlnThrAsnThr--------------- 39
                             :::      |||:::   :::|||:::
Db     268 ACTAATCAAAGTGGACCCAACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAAGAG 327

Qy      40 ---------TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56
                    |||::::::                |||||    ||||||||
Db     328 ATGGCCATATACGCCTAC---------------TATTACACCGGGATTGGTGCTGGTGTG 372

Qy      57 LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys 76
           |||||         |||||||||:::||||||||     :::   |||   |||||
Db     373 CTCATAGTTGCCTACATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGAAGACAGATACAC 432

Qy      77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96
           :::||||||::::::|||||||||:::::::::   ||||:::||||||||||||||||
Db     433 AAGATTAGGCAGAAGTTTTTCCATGCTATAATGAATCAGGAGATAGGCTGGTTTGATGTG 492

Qy      97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
              |||:::||||||||||||||||||:::|||   |||:::     ||||||:::|||||
Db     493 CATGATGTTGGGGAGCTCAACACCCGGCTCACAGATGATGTCTCCAAAATTAATGACGGA 552

Qy     116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
           ||||||||||||||    :::   ||||||:::::::::||||||         |||
Db     553 ATTGGTGACAAAATTGGGATGTTTTTTCAGTCCATAACCACATTTTTAGCCGGTTTTATC 612

Qy     136 ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
           :::|||   :::   ||||||||||||||||||||   |||:::   ||||||||||||
Db     613 ATAGGATTTATAAGTGGTTGGAAGCTAACCCTTGTCATTTTGGCTGTCAGCCCTCTTATT 672

Qy     156 MetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
                 |||:::|||      ::::::::::::   |||    |||:::|||||||||
Db     673 GGATTGTCATCTGCTTTGTGGGCAAAGGTATTGACTTCATTTACTAATAAGGAACTCCAG 732

Qy     176 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIle 195
           ||||||:::|||||||||||||||||||||||||||||||||::::::||||||||||||
Db     733 GCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT 792

Qy     196 AlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLys 215
           ||||||      |||:::|||||||||:::||||||      :::||||||::::::||||||
Db     793 GCCTTTGGAGGACAACAGAAGGAACTTGAAAGGTACAATAAAAATTTAGAAGAAGCTAAA 852

Qy     216 AspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePhe 235
           :::   |||||||||:::   |||   :::   :::|||:::|||        |||
Db     853 AATGTTGGCATAAAGAAAGCTATCACAGCCAGCATTTCGATAGGCATTGCCTACCTGTTG 912

Qy     236 MetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGlu 255
           :::       :::|||   ||||||||||||||||||||||||:::|||:::      |||
Db     913 GTCTATGCATCATATGCACTGGCATTCTGGTATGGGACATCCTTGGTCCTCTCAAATGAA 972

Qy     256 ProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyr 275
```

```
Db     973 ------TATTCTATTGGAGAAGTGCTTACTGTCTTCTTCTCTATTTGTTGGGGACTTTT 1026

Qy     276 CysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPhe 295
Db    1027 AGTATTGGACACTTGGCCCCAAACATAGAAGCCTTTGCAAACGCACGAGGGGCAGCCTTT 1086

Qy     296 HisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyr 315
Db    1087 GAAATCTTCAAGATAATTGATAACGAGCCAAGCATTGACAGCTTCTCAACAAAGGGCTAC 1146

Qy     316 LysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSer 335
Db    1147 AAACCAGACAGTATAATGGGAAACTTAGAGTTTAAAAATGTTCACTTCAACTACCCATCG 1206

Qy     336 ArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrVal 355
Db    1207 AGAAGCGAAGTTCAGATCTTGAAGGGCCTCAATCTGAAGGTGAAGAGCGGACAGACGGTG 1266

Qy     356 AlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeu 375
Db    1267 GCCTTGGTTGGCAACAGTGGCTGTGGAAAAAGCACAACTGTCCAGCTGATGCAGAGGCTC 1326

Qy     376 TyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVal 395
Db    1327 TACGACCCCCTGGAGGGCGTGGTCAGTATCGACGGACAAGACATCAGAACCATCAATGTG 1386

Qy     396 ArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThr 415
Db    1387 AGGTATCTGAGGGAGATCATTGGTGTGGTGAGTCAGGAACCTGTGCTGTTTGCCACCACG 1446

Qy     416 IleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAla 435
Db    1447 ATCGCCGAGAACATTCGCTATGGCCGAGAAGATGTCACCATGGATGAGATTGAGAAAGCT 1506

Qy     436 AlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeu 455
Db    1507 GTCAAGGAAGCCAATGCCTATGACTTCATCATGAAACTGCCCCACCAATTTGACACCCTG 1566

Qy     456 ValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArg 475
Db    1567 GTTGGTGAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGAATCGCCATTGCCCGG 1626

Qy     476 AlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSer 495
Db    1627 GCCCTGGTCCGCAATCCCAAGATCCTTTTGTTGGACGAGGCCACCTCAGCCCTGGATACA 1686

Qy     496 GluSerLysSerAlaValGlnAlaAlaLeuGluLys 507
Db    1687 GAAAGTGAAGCTGTGGTGCAGGCCGCACTGGATAAG 1722
```

RESULT 7
US-09-866-866A-1
; Sequence 1, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 3860
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-866-866A-1

```
Alignment Scores:
Pred. No.:               8.59e-159     Length:          3860
Score:                   1579.00       Matches:         306
Percent Similarity:      77.04%        Conservative:    100
Best Local Similarity:   58.06%        Mismatches:      99
Query Match:             60.92%        Indels:          22
DB:                      10            Gaps:            5
```

US-09-873-409-8 (1-514) x US-09-866-866A-1 (1-3860)

```
Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db     151 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210

Qy      21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
Db     211 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 270

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db     271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db     331 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 390

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db     391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db     451 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 510

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db     511 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGGTATTGGTGACAAAATT 570

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db     571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 630

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db     631 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db     691 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db     751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810

Qy     201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db     811 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 870

Qy     221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db     871 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 930

Qy     241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db     931 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 984

Qy     261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db     985 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1044
```

```
Qy     281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
               |||      |||    ||||||   |||||||||||||||:::   ||||||::::::
Db    1045 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104

Qy     301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
           ||||||   ||||||||||||||||:::::::|||   :::|||:::||||||::::::|||
Db    1105 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1164

Qy     321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
           :::|||   :::||||||:::||||||   |||:::|||||||||||         :::|||
Db    1165 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1224

Qy     341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
           ||||||||||||||||||||||||:::::::::||||||:::||||||||||||||||||
Db    1225 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1284

Qy     361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
           :::|||   |||||||||||||   |||||||||:::||||||||||||||||||   :::
Db    1285 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1344

Qy     381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
           |||   :::   ||||||      |||||||||      :::|||||||||   |||:::
Db    1345 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1404

Qy     401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
              |||||||||||||||||||||||||||||||||   ||||||||:::    ||||||
Db    1405 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1464

Qy     421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
           :::|||||||||::::::||||||   :::|||:::|||:::|||   :::|||||||||
Db    1465 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1524

Qy     441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
           ||||||||||||||||||:::   |||:::||||||:::||||||||||||||||:::|||
Db    1525 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1584

Qy     461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
           ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1585 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1644

Qy     481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
           ||||||||||:::|||||||||||||||||||||||||||||:::|||||||::::::
Db    1645 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1704

Qy     501 ValGlnAlaAlaLeuGluLys 507
           ||||||   ||||||:::|||
Db    1705 GTTCAGGTGGCTCTGGATAAG 1725
```

RESULT 8
US-10-044-671-1
; Sequence 1, Application US/10044671
; Patent No. US20020177147A1
; GENERAL INFORMATION:
; APPLICANT: Washington State University Research Foundation
; APPLICANT: Mealey, Katrina
; APPLICANT: Bentjen, Steven
; TITLE OF INVENTION: MDR1 VARIANTS AND METHODS FOR THEIR USE
; FILE REFERENCE: 4630-61733
; CURRENT APPLICATION NUMBER: US/10/044,671
; CURRENT FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/261,578
; PRIOR FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/314,829
; PRIOR FILING DATE: 2001-08-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
;   LENGTH: 4317
;   TYPE: DNA
;   ORGANISM: Canis familiaris
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (70)..(3912)
;   OTHER INFORMATION:
US-10-044-671-1

Alignment Scores:

Pred. No.:	2.13e-158	Length:	4317
Score:	1576.00	Matches:	302
Percent Similarity:	75.61%	Conservative:	98
Best Local Similarity:	57.09%	Mismatches:	105
Query Match:	60.80%	Indels:	24
DB:	9	Gaps:	4

US-09-873-409-8 (1-514) x US-10-044-671-1 (1-4317)

```
Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||::::::|||   :::|||::::::::::::|||||||||   ||||||||||||   |||
Db     217 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 276

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
           |||   |||   |||:::|||:::         :::         :::
Db     277 GTTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 336

Qy      35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                                :::      ||||||      |||
Db     337 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 396

Qy      45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                          ||||||      |||||||||         |||:::
Db     397 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 456

Qy      60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
              |||||||||:::|||   |||   :::   |||   ||||||      :::||||||
Db     457 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 516

Qy      80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
           |||||||||||||||:::::::::   |||:::|||||||||||||||||||   |||:::
Db     517 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 576

Qy     100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
           ||||||||||||||||||:::|||   |||:::   ||||||::::::||||||||||||
Db     577 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 636

Qy     119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
           |||:::   :::   ||||||:::::::::||||||      |||      ||||||
Db     637 AAAGTTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 696

Qy     139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
                 |||||||||||||||||||   |||:::   ||||||::::::      |||
Db     697 ACACCTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 756

Qy     159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
           ||||||      ::::::::::::   |||   |||   |||||||||   ||||||:::
Db     757 GCCGCCATCTGGGCAAAGATACTATCTTCATTACTGATAAAGAACTCTTGGCCTATGCA 816

Qy     179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
           |||||||||||||||||||||||||||||::::::|||||||||||||||||||||||||
Db     817 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 876

Qy     199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
              |||:::|||||||||:::||||||   :::||||||::::::||||||      |||
Db     877 GGACAAAAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGG 936

Qy     219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238
           ||||||:::   |||   :::   :::|||:::||||||      :::      :::
Db     937 ATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCA 996

Qy     239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
           :::|||   ||||||||||||||||||||||||:::|||:::      |||      |||
Db     997 TCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TAT 1050

Qy     259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
           |||||||||   ||||||   ||||||||||||:::         ::::::   ||||||
Db    1051 ACTATTGGACAGGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGA 1110
```

```
Qy     279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
              |||  |||     |||   ||||||  ||||||||||||||:::   ||||||
Db    1111 CAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTC 1170

Qy     299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
           ::::::||||||  |||||||||||||||::::::|||    :::|||:::||||||:::
Db    1171 AAGATAATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGAT 1230

Qy     319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
           :::|||:::|||   :::||||||||||||||||    |||:::||||||||||||
Db    1231 AATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAA 1290

Qy     339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
           :::||||||||||||||||||||||||:::::::::||||||:::||||||||||||||||
Db    1291 GTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTT 1350

Qy     359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
           |||   :::|||   ||||||||||||   |||||||||:::||||||||||||||||||
Db    1351 GGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCC 1410

Qy     379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
              ||||||   :::    :::|||    |||||||||     :::||||||||||||
Db    1411 ACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTT 1470

Qy     399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
           |||:::      ||||||||||||||||||||||||||||||   |||||||||:::
Db    1471 CGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAA 1530

Qy     419 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
           ||||||:::|||||||||:::::::||||||    :::|||:::|||:::|||    :::|||
Db    1531 AACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAA 1590

Qy     439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
           |||||||||||||||||||||||||||:::    ||||||||||||:::||||||||||||||||
Db    1591 GCCAATGCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAG 1650

Qy     459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
           :::||||||:::::::||||||||||||||||||||||||||||||||||||||||||||
Db    1651 AGAGGGGCCCGGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTT 1710

Qy     479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
           ||||||||||||||||||:::||||||||||||||||||||||||||||:::||||||:::
Db    1711 CGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAA 1770

Qy     499 SerAlaValGlnAlaAlaLeuGluLys 507
           :::   ||||||   ||||||:::|||
Db    1771 GCAGTGGTTCAGGTGGCCCTGGATAAG 1797
```

RESULT 9
US-09-917-800A-1424
; Sequence 1424, Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44921-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: US 60/290,029
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/290,645
; PRIOR FILING DATE: 2001-05-15
; PRIOR APPLICATION NUMBER: US 60/292,336
; PRIOR FILING DATE: 2001-05-22
; PRIOR APPLICATION NUMBER: US 60/295,798
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/297,457
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 60/298,884
; PRIOR FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US 60/303,459
; PRIOR FILING DATE: 2001-07-09
; NUMBER OF SEQ ID NOS: 1740
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1424
;   LENGTH: 4254
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
;   FEATURE:
;   OTHER INFORMATION: Genbank Accession No. US20020119462A1 M81855
US-09-917-800A-1424

```
Alignment Scores:
Pred. No.:                6.32e-158      Length:        4254
Score:                    1571.50        Matches:       307
Percent Similarity:       75.95%         Conservative:  94
Best Local Similarity:    58.14%         Mismatches:    104
Query Match:              60.63%         Indels:        23
DB:                       10             Gaps:          4
```

US-09-873-409-8 (1-514) x US-09-917-800A-1424 (1-4254)

```
Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||   ||||||   ||||||::::::::::::|||         |||||||||:::   |||
Db     247 ATGGCTCTGGGAACTCTCGCTGCTATCATCCACGGAACCCTGCTTCCCCTCCTGATGCTG 306

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
           |||   |||   |||:::|||:::
Db     307 GTGTTCGGATACATGACAGATAGTTTTACCCCAAGCAGAGACCCGCATTCTGACCGAGCG 366

Qy      31 ---------SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu------ 45
                             :::      |||:::|||   |||         |||
Db     367 ATTACTAATCAAAGTGAAATCAACAGTACACATACCGTCAGCGACACGAGTCTGGAGGAG 426

Qy      46 ---------------ThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
                                 ||||||   ||||||||||      ||||||
Db     427 GACATGGCCATGTATGCCTACTATTACACGGGCATTGGTGCCGGTGTGCTCATCGTTGCC 486

Qy      61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
           |||||||||:::||||||||||   :::   |||   ||||||      :::||||||:::
Db     487 TACATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAG 546

Qy      81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
           :::||||||||||::::::::::   |||:::|||||||||||||||||      |||   |||
Db     547 AAGTTTTTCCATGCCATCATGAATCAGGAGATAGGCTGGTTTGACGTGAATGACGCTGGG 606

Qy     101 GluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLys 119
           |||||||||||||||:::|||   |||:::   ||||||:::||||||||||||||||||
Db     607 GAGCTCAACACCCGTCTCACAGATGACGTCTCCAAAATTAATGACGGAATTGGTGACAAA 666

Qy     120 IleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuVal 139
           :::   :::   ||||||::::::::::|||||||||   |||         :::|||   :::
Db     667 CTTGGAATGTTCTTTCAGTCCATAACGACATTTTCAGCCGGTTTTATAATAGGATTTATA 726

Qy     140 LysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAla 159
              |||||||||||||||||||||   |||:::   ||||||||||||      |||:::
Db     727 AGTGGTTGGAAGCTAACCCTTGTAATTTTGGCCGTCAGCCCTCTTATTGGGTTGTCATCT 786

Qy     160 AlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLys 179
           |||      ::::::::::::   |||   |||:::|||||||||   ||||||:::|||
Db     787 GCCATGTGGGCAAAGGTACTGACTTCATTTACTAATAAGGAACTCCAGGCTTATGCGAAA 846

Qy     180 AlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAla 199
           |||||||||||||||||||||||||||:::::::|||||||||||||||||||||
Db     847 GCTGGAGCAGTTGCCGAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCGTTTGGAGGA 906

Qy     200 GlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIle 219
```

```
              |||:::||||||||:::||||||   :::||||||::::::||||||      ||||||
Db       907 CAAAAGAAGGAACTTGAAAGGTACAATAAAAATTTAGAAGAAGCTAAAAGAGTTGGCATA 966

Qy       220 LysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThr 239
             |||:::   |||   :::   :::|||:::|||       |||       :::       :::
Db       967 AAGAAAGCCATCACGGCCAACATTTCCATAGGTATTGCCTACCTGTTGGTCTATGCGTCT 1026

Qy       240 TyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThr 259
             |||   |||||||||||||||||||||||||||||||||:::|||:::   |||   |||:::
Db      1027 TATGCACTGGCATTCTGGTATGGGACCTCCTTGGTCCTCTCAAATGAA------TATTCT 1080

Qy       260 IleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAla 279
             ||||||   ||||||   ||||||||||||::::::       ::::::   ||||||
Db      1081 ATTGGACAAGTGCTTACCGTCTTCTTCTCTATTTATTGGGGACTTTCAGTATTGGACAT 1140

Qy       280 AlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGln 299
                      |||:::   |||   ||||||   |||||||||||||||:::   ||||||:::
Db      1141 TTAGCCCCAAACATAGAAGCCTTTGCAAATGCAAGAGGGGCAGCCTATGAAATCTTCAAG 1200

Qy       300 ValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSer 319
             :::||||||   :::||||||||||||:::|||||||||      |||:::||||||:::|||
Db      1201 ATAATTGATAATGAGCCAAGCATCGACAGCTTCTCAACCAAGGGACACAAACCAGACAGT 1260

Qy       320 IleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIle 339
             |||   |||   :::||||||||||||      ||||||||||||||||||         :::
Db      1261 ATAATGGGAAATTTGGAATTTAAAAATGTTTACTTCAACTACCCATCACGAAGTGAAGTT 1320

Qy       340 LysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGly 359
             |||||||||||||||||||||||::::::|||||||:::||||||||||||||||||||
Db      1321 AAGATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTGGTTGGC 1380

Qy       360 LeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAsp 379
                :::|||   |||||||||||   |||||||||||||||||||||||||||
Db      1381 AACAGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCATA 1440

Qy       380 AspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArg 399
             :::|||   :::   :::|||         ||||||||   :::|||||||||:::   |||
Db      1441 GAGGGCGAGGTCAGTATTGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGG 1500

Qy       400 AspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsn 419
             :::   ||||||||||||||||||||||||||||||   |||||||||:::      |||
Db      1501 GAAATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCCACCACGATTGCCGAAAAC 1560

Qy       420 IleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAla 439
             |||:::||||||||:::::::|||:::|||:::|||      :::||||||
Db      1561 ATTCGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCC 1620

Qy       440 AsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLys 459
             ||||||||||||||||||||:::      |||:::||||||:::||||||||||||||:::
Db      1621 AATGCCTATGACTTCATCATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAGAGA 1680

Qy       460 GlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArg 479
             |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db      1681 GGGGCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGC 1740

Qy       480 AsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSer 499
             ||||||||||||||:::|||||||||||||||||||||||||||:::||||||::::::
Db      1741 AACCCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCC 1800

Qy       500 AlaValGlnAlaAlaLeuGluLys 507
                |||||||||||||||:::|||
Db      1801 GTGGTTCAGGCCGCTCTGGATAAG 1824
```

RESULT 10
US-09-866-866A-3
; Sequence 3, Application US/09866866A
; Patent No. US20020102244A1
; GENERAL INFORMATION:
; APPLICANT: Sorrentino, Brian
; APPLICANT: Schuetz, John
; TITLE OF INVENTION: A Method of Identifying and/or Isolating Stem Cells
; FILE REFERENCE: 1340-1-021CIP2
; CURRENT APPLICATION NUMBER: US/09/866,866A
; CURRENT FILING DATE: 2001-08-30
; PRIOR APPLICATION NUMBER: 09/584,586
; PRIOR FILING DATE: 2000-05-31
; PRIOR APPLICATION NUMBER: PCT/US99/11825
; PRIOR FILING DATE: 1999-05-27
; PRIOR APPLICATION NUMBER: 60/086,988
; PRIOR FILING DATE: 1998-05-28
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
;   LENGTH: 3860
;   TYPE: DNA
;   ORGANISM: homo sapiens
US-09-866-866A-3

```
Alignment Scores:
Pred. No.:              7.89e-158       Length:         3860
Score:                  1570.00         Matches:        305
Percent Similarity:     76.85%          Conservative:   100
Best Local Similarity:  57.87%          Mismatches:     100
Query Match:            60.57%          Indels:         22
DB:                     10              Gaps:           5

US-09-873-409-8 (1-514) x US-09-866-866A-3 (1-3860)

Qy         1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
             |||::::::|||   ||||||::::::::::::::||||||   |||||||||||   |||
Db       151 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210

Qy        21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
             |||   ||||||||::::|||                          :::|||:::|||
Db       211 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 270

Qy        32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                      ::::::         |||         ||||||
Db       271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330

Qy        44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                   |||         ||||||   |||||||||         |||:::      |||
Db       331 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 390

Qy        62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
             ||||||:::|||   |||   :::   |||   ||||||      :::||||||||||||
Db       391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450

Qy        82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
             ||||||||:::::::::      |||:::||||||||||||||||      |||:::|||||
Db       451 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 510

Qy       102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
             |||||||||||:::|||   |||:::   ::::::   ||||||||||||||||
Db       511 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 570

Qy       121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
                :::   ||||||:::|||:::||||||         |||      ||||||      :::
Db       571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 630

Qy       141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
             ||||||||||||||||||||   |||:::   ||||||::::::      |||||||||
Db       631 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

Qy       161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
                ::::::::::::      |||   |||   |||||||||   ||||||:::||||||
Db       691 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy       181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
             |||||||||||||||||||||||||||||||||||||||||||||||||         |||
Db       751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810

Qy       201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
             :::||||||||||:::||||||   :::||||||::::::||||||      |||||||||
Db       811 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 870
```

```
Qy     221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
           :::   |||   :::   :::|||:::|||||   :::       :::      :::|||
Db     871 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 930

Qy     241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
           ||||||||||||||||||||||||:::|||:::|||:::||||||      |||:::|||
Db     931 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 984

Qy     261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
           |||   ||||||   ||||||||||||||||:::        ::::::   :::|||   |||
Db     985 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTAGTGTTGGACAGGCA 1044

Qy     281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
              |||      |||   ||||||   ||||||||||||||:::   ||||||::::::
Db    1045 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104

Qy     301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
           ||||||   ||||||||||||||::::::|||   :::|||:::||||||::::::|||
Db    1105 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1164

Qy     321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
           :::|||   :::||||||:::||||||   |||:::||||||||||||      :::|||
Db    1165 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1224

Qy     341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
           ||||||||||||||||||||||||:::::::::||||||:::||||||||||||||||||
Db    1225 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1284

Qy     361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
           :::|||   ||||||||||||   ||||||||:::||||||||||||||||||   :::
Db    1285 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1344

Qy     381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
           |||   :::   ||||||      ||||||||||   :::||||||||||      |||:::
Db    1345 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1404

Qy     401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
              |||||||||||||||||||||||||||||||||||   |||||||:::   ||||||
Db    1405 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1464

Qy     421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
           :::||||||||||:::::::||||||   :::|||:::|||:::|||   :::|||||||||
Db    1465 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1524

Qy     441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
           ||||||||||||||||||:::      |||:::||||||:::||||||||||||:::|||
Db    1525 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1584

Qy     461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
           ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1585 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1644

Qy     481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
           ||||||||||:::||||||||||||||||||||||||||||:::||||||::::::
Db    1645 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1704

Qy     501 ValGlnAlaAlaLeuGluLys 507
           ||||||   ||||||:::|||
Db    1705 GTTCAGGTGGCTCTGGATAAG 1725
```

RESULT 11
US-09-306-417-1
; Sequence 1, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
; APPLICANT: Heinrich-Pette-Institut
; TITLE OF INVENTION: Retroviral Gene Transfer Vectors
; FILE REFERENCE: P50491
; CURRENT APPLICATION NUMBER: US/09/306,417
; CURRENT FILING DATE: 1999-05-06
; EARLIER APPLICATION NUMBER: DE 198 22 115
; EARLIER FILING DATE: 1998-05-08
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 8630
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: proviral
; OTHER INFORMATION: plasmid DNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(160)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (161)..(677)
; OTHER INFORMATION: 5'-LTR
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (532)..(1219)
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: (1220)..(5062)
; OTHER INFORMATION: m4 mdr-1 cDNA
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5215)..(5774)
; OTHER INFORMATION: 3'-LTR
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (5775)..(8630)
; OTHER INFORMATION: plasmid backbone (pUC)
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(8630)
; OTHER INFORMATION: retroviral expression vector SFbeta71m4
US-09-306-417-1

Alignment Scores:

Pred. No.:	2.68e-157	Length:	8630
Score:	1570.00	Matches:	305
Percent Similarity:	76.85%	Conservative:	100
Best Local Similarity:	57.87%	Mismatches:	100
Query Match:	60.57%	Indels:	22
DB:	10	Gaps:	5

US-09-873-409-8 (1-514) x US-09-306-417-1 (1-8630)

```
Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||:::::::|||   ||||||::::::::::::||||||   ||||||||||||   |||
Db    1370 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 1429

Qy      21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
           |||   ||||||||:::|||                           :::|||:::|||
Db    1430 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 1489

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                    ::::::      |||      ||||||
Db    1490 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 1549

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                 |||      ||||||   |||||||||      |||:::       |||
Db    1550 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 1609

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
           ||||||:::|||   |||   :::   |||   ||||||      :::||||||||||||
Db    1610 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 1669

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
           ||||||||:::::::::   |||:::||||||||||||||      |||:::||||||
Db    1670 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 1729

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
           ||||||||||||:::|||   |||:::   ||||||::::::   ||||||||||||||
```

```
Db    1730 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 1789

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    1790 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1849

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    1850 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1909

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    1910 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1969

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    1970 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 2029

Qy     201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db    2030 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 2089

Qy     221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db    2090 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 2149

Qy     241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db    2150 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 2203

Qy     261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db    2204 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 2263

Qy     281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db    2264 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 2323

Qy     301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db    2324 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 2383

Qy     321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db    2384 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 2443

Qy     341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db    2444 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 2503

Qy     361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db    2504 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 2563

Qy     381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db    2564 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 2623

Qy     401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db    2624 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 2683

Qy     421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db    2684 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 2743

Qy     441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db    2744 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2803

Qy     461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db    2804 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2863

Qy     481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db    2864 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2923

Qy     501 ValGlnAlaAlaLeuGluLys 507
Db    2924 GTTCAGGTGGCTCTGGATAAG 2944

RESULT 12
US-09-306-417-2
; Sequence 2, Application US/09306417
; Patent No. US20020103144A1
; GENERAL INFORMATION:
;  APPLICANT: Heinrich-Pette-Institut
;  TITLE OF INVENTION: Retroviral Gene Transfer Vectors
;  FILE REFERENCE: P50491
;  CURRENT APPLICATION NUMBER: US/09/306,417
;  CURRENT FILING DATE: 1999-05-06
;  EARLIER APPLICATION NUMBER: DE 198 22 115
;  EARLIER FILING DATE: 1998-05-08
;  NUMBER OF SEQ ID NOS: 17
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 8630
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: proviral
;   OTHER INFORMATION: plasmid DNA
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(8630)
;   OTHER INFORMATION: retroviral expression vector SFbeta91mSA1
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)..(160)
;   OTHER INFORMATION: plasmid backbone (pUC)
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (161)..(677)
;   OTHER INFORMATION: 5'-LTR
;   FEATURE:
;   NAME/KEY: 5'UTR
;   LOCATION: (532)..(1219)
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: (1220)..(5062)
;   OTHER INFORMATION: mSA1 mdr1 cDNA
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5215)..(5774)
;   OTHER INFORMATION: 3'-LTR
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (5775)..(8630)
;   OTHER INFORMATION: plasmid backbone (pUC)
US-09-306-417-2

Alignment Scores:
Pred. No.:              2.68e-157     Length:         8630
Score:                  1570.00       Matches:        305
Percent Similarity:     76.85%        Conservative:   100
Best Local Similarity:  57.87%        Mismatches:     100
Query Match:            60.57%        Indels:         22
DB:                     10            Gaps:           5

US-09-873-409-8 (1-514) x US-09-306-417-2 (1-8630)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    1370 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 1429
```

```
Qy     21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db   1430 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 1489

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db   1490 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 1549

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db   1550 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 1609

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db   1610 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 1669

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db   1670 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 1729

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db   1730 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 1789

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db   1790 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1849

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db   1850 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1909

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db   1910 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1969

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db   1970 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 2029

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db   2030 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 2089

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db   2090 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 2149

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db   2150 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 2203

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db   2204 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 2263

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db   2264 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 2323

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db   2324 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 2383

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db   2384 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 2443

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   2444 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 2503

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   2504 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 2563

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   2564 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 2623

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   2624 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 2683

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   2684 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 2743

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   2744 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2803

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   2804 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2863

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   2864 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2923

Qy    501 ValGlnAlaAlaLeuGluLys 507
Db   2924 GTTCAGGTGGCTCTGGATAAG 2944
```

RESULT 13
US-09-769-097-3
; Sequence 3, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: ENCODING RAT MDR1A AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP-50009-C2
; CURRENT APPLICATION NUMBER: US/09/769,097
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 3
;   LENGTH: 4425
;   TYPE: DNA
;   ORGANISM: RATTUS RATTUS
US-09-769-097-3

Alignment Scores:			
Pred. No.:	2.3e-157	Length:	4425
Score:	1566.50	Matches:	300
Percent Similarity:	76.82%	Conservative:	101
Best Local Similarity:	57.47%	Mismatches:	102
Query Match:	60.44%	Indels:	19
DB:	10	Gaps:	4

US-09-873-409-8 (1-514) x US-09-769-097-3 (1-4425)

```
Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    499 ATGCTGCTGGGAACTCTGGCGGCCATTATCCATGGAATTGCGCTCCCACTTATGATGCTG 558
```

```
Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db    559 GTCTTTGGAGACATGACAGATAGCTTTGCAAAT------GTAGGAAACAACCGTAGTATG 612

Qy     41 SerPhePheArgLeuThr---------------------------------------- 46
Db    613 AGTTTCTACAATGCTACAGACATATATGCCAAGCTGGAGGACGAAATGACCACGTACGCC 672

Qy     47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
Db    673 TACTATTACACGGGCATTGGTGCCGGTGTGCTCATCGTTGCCTACATCCAGGTTTCCACT 732

Qy     67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
Db    733 TGGTGCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAAGTTTTTCCATGCCATC 792

Qy     87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
Db    793 ATGAATCAGGAGATAGGCTGGTTTGACGTGCATGACGTTGGGGAGCTCAACACCCGGCTC 852

Qy    107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
Db    853 ACAGATGACGTCTCCAAAATTAATGAAGGAATTGGTGACAAAATTGGAATGTTCTTTCAG 912

Qy    126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
Db    913 GCAATGGCAACATTTTTGGTGGTTTTATAATAGGATTTACTCGCGGCTGGAAGCTAACT 972

Qy    146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
Db    973 CTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTGTCAGCTGGTATTTGGGCAAAGATA 1032

Qy    166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
Db   1033 TTGTCTTCATTTACTGATAAGGAACTCCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAA 1092

Qy    186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
Db   1093 GAAGTCTTAGCAGCCATCAGAACTGTGATTGCCTTTGGAGGACAAAAGAAGGAACTTGAA 1152

Qy    206 ArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSer 225
Db   1153 AGGTACAATAACAATTTGGAAGAAGCTAAAAGGCTTGGGATAAAGAAAGCTATCACGGCC 1212

Qy    226 LysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrp 245
Db   1213 AACATTTCCATGGGTGCAGCTTTTCTGCTTATCTATGCATCATATGCTCTGGCATTCTGG 1272

Qy    246 TyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAla 265
Db   1273 TATGGGACTTCCTTGGTCATCTCAAAAGAA------TACACTATTGGACAAGTGCTCACT 1326

Qy    266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285
Db   1327 GTCTTTTTTTCTGTATTAATTGGAGCATTCAGTGTTGGGCAGGCATCTCCAAATATTGAA 1386

Qy    286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305
Db   1387 GCCTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTCAGTATAATTGATAATAAGCCC 1446

Qy    306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
Db   1447 AGTATAGACAGCTTCTCAAAGAGTGGGCACAAACCCGACAACATACAAGGAAATTTGGAA 1506

Qy    326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
Db   1507 TTCAAAAATATTCACTTCAGTTACCCGTCTCGAAAAGACGTTCAGATCTTGAAGGGCCTC 1566

Qy    346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
Db   1567 AACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTGGTTGGCAACAGTGGCTGTGGGAAA 1626

Qy    366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
Db   1627 AGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCATAGAGGGCGAGGTCAGTATC 1686

Qy    386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
Db   1687 GACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTGGTG 1746

Qy    406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
Db   1747 AGTCAGGAACCCGTGCTGTTTGCCACCACAATTGCCGAAAACATTCGCTATGGCCGAGAA 1806

Qy    426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
Db   1807 AACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAATGCCTATGATTTCATC 1866

Qy    446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465
Db   1867 ATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGG 1926

Qy    466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
Db   1927 GGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAACCCCAAGATCCTTTTG 1986

Qy    486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeu 505
Db   1987 TTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTGGTTCAGGCCGCTCTG 2046

Qy    506 GluLys 507
Db   2047 GATAAG 2052
```

RESULT 14
US-09-769-097-1
; Sequence 1, Application US/09769097
; Patent No. US20020055128A1
; GENERAL INFORMATION:
; APPLICANT: Kimberly Anne Brun
; APPLICANT: Richard James Chenery
; APPLICANT: Harma Ellens
; APPLICANT: John Anthony Feild
; APPLICANT: Lin Yue
; TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
; TITLE OF INVENTION: ENCODING RAT MDR1A AND SCREENING METHODS THEREOF
; FILE REFERENCE: GP-50009-C2
; CURRENT APPLICATION NUMBER: US/09/769,097
; CURRENT FILING DATE: 2001-01-25
; PRIOR APPLICATION NUMBER: 09/208,809
; PRIOR FILING DATE: 1998-12-09
; PRIOR APPLICATION NUMBER: 09/156,800
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: US99/20770
; PRIOR FILING DATE: 1999-09-10
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 4369
; TYPE: DNA
; ORGANISM: RATTUS RATTUS
US-09-769-097-1

Alignment Scores:

Pred. No.:	3.7e-157	Length:	4369
Score:	1564.50	Matches:	300
Percent Similarity:	76.82%	Conservative:	101
Best Local Similarity:	57.47%	Mismatches:	102
Query Match:	60.36%	Indels:	19
DB:	10	Gaps:	4

US-09-873-409-8 (1-514) x US-09-769-097-1 (1-4369)

```
Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
```

```
Db     499 ATGCTGCTGGGAACTCTGGCGGCCATTATCCATGGAATTGCGCTCCCACTTATGATGCTG 558

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db     559 GTCTTTGGAGACATGACAGATAGCTTTGCAAAT------GTAGGAAACAACCGTAGTATG 612

Qy      41 SerPhePheArgLeuThr------------------------------------------ 46
Db     613 AGTTTCTACAATGCTACAGACATATATGCCAAGCTGGAGGACGAAATGGCCACGTACGCC 672

Qy      47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
Db     673 TACTATTACACGGGCAATGGTGCCGGTGTGCTCATCGTTGCCTACATCCAGGTTTCACTT 732

Qy      67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
Db     733 TGGTGCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAAGTTTTTCCATGCCATC 792

Qy      87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
Db     793 ATGAATCAGGAGATAGGCTGGTTTGACGTGCATGACGTTGGGGAGCTCAACACCCGGCTC 852

Qy     107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
Db     853 ACAGATGACGTCTCCAAAATTAATGAAGGAATTGGTGACAAAATTGGAATGTTCTTTCAG 912

Qy     126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
Db     913 GCAATGGCAACATTTTTTGGTGGTTTTATAATAGGATTTACTCGCGGCTGGAAGCTAACT 972

Qy     146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
Db     973 CTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTGTCAGCTGGTATTTGGGCAAAGATA 1032

Qy     166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
Db    1033 TTGTCTTCATTTACTGATAAGGAACTCCAGGCTTATGCAAAAGCTGGAGCAGTTGCTGAA 1092

Qy     186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
Db    1093 GAAGTCTTAGCAGCCATCAGAACTGTGATTGCCTTTGGAGGACAAAAGAAGGAACTTGAA 1152

Qy     206 ArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSer 225
Db    1153 AGGTACAATAACAATTTGGAAGAAGCTAAAAGGCTTGGGATAAAGAAAGCTATCACGGCC 1212

Qy     226 LysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrp 245
Db    1213 AACATTTCCATGGGTGCAGCTTTTCTGCTTATCTATGCATCATATGCTCTGGCATTCTGG 1272

Qy     246 TyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAla 265
Db    1273 TATGGGACTTCCTTGGTCATCTCAAAAGAA------TACACTATTGGACAAGTGCTCACT 1326

Qy     266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285
Db    1327 GTCTTTTTTTCTGTATTAATTGGAGCATTCAGTGTTGGGCAGGCATCTCCAAATATTGAA 1386

Qy     286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305
Db    1387 GCCTTCGCCAATGCTAGAGGAGCAGCTTATGAAGTCTTCAGTATAATTGATAATAAGCCC 1446

Qy     306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
Db    1447 AGTATAGACAGCTTCTCAAAGAGTGGGCACAAACCCGACAACATACAAGGAAATTTGGAA 1506

Qy     326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
Db    1507 TTCAAAAATATTCACTTCAGTTACCCGTCTCGAAAAGACGTTCAGATCTTGAAGGGCCTC 1566

Qy     346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
Db    1567 AACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTGGTTGGCAACAGTGGCTGTGGGAAA 1626

Qy     366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
Db    1627 AGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCATAGAGGGCGAGGTCAGTATC 1686

Qy     386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
Db    1687 GACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAAATCATTGGGGTGGTG 1746

Qy     406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
Db    1747 AGTCAGGAACCCGTGCTGTTTGCCACCACAATTGCCGAAAACATTCGCTATGGCCGAGAA 1806

Qy     426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
Db    1807 AACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAATGCCTATGATTTCATC 1866

Qy     446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465
Db    1867 ATGAAACTGCCCCACAAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGG 1926

Qy     466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
Db    1927 GGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAACCCCAAGATCCTTTTG 1986

Qy     486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeu 505
Db    1987 TTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTGGTTCAGGCCGCTCTG 2046

Qy     506 GluLys 507
Db    2047 GATAAG 2052
```

RESULT 15
US-10-072-621-2
; Sequence 2, Application US/10072621
; Patent No. US20020169137A1
; GENERAL INFORMATION:
; APPLICANT: Reiner, Peter B.
; APPLICANT: Connop, Bruce P.
; APPLICANT: Pollard, Michelle
; TITLE OF INVENTION: REGULATION OF AMYLOID PRECURSOR PROTEIN EXPRESSION
; TITLE OF INVENTION: BY MODIFICATION OF ABC TRANSPORTER EXPRESSION OR ACTIVITY
; FILE REFERENCE: 100103.402
; CURRENT APPLICATION NUMBER: US/10/072,621
; CURRENT FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 4643
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-072-621-2

Alignment Scores:

Pred. No.:	6.64e-157	Length:	4643
Score:	1562.50	Matches:	305
Percent Similarity:	76.85%	Conservative:	100
Best Local Similarity:	57.87%	Mismatches:	99
Query Match:	60.28%	Indels:	23
DB:	9	Gaps:	6

US-09-873-409-8 (1-514) x US-10-072-621-2 (1-4643)

```
Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db     575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy      21 ValLeuGlyGluMetSerAsp-------------------------AsnLeuIleSer 31
Db     635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
```

```
Db      695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy       44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db      755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy       62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db      815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy       82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db      875 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy      102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db      935 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAAATT 994

Qy      121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db      995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy      141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db     1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy      161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db     1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy      181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db     1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy      201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db     1235 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1294

Qy      221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db     1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1354

Qy      241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db     1355 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1408

Qy      261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db     1409 GGACAAGTACTCACTGTA---TTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1465

Qy      281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db     1466 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1525

Qy      301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db     1526 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1585

Qy      321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db     1586 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1645

Qy      341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db     1646 ATCTTGAAGGGTCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1705

Qy      361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db     1706 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1765

Qy      381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400

Db     1766 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1825

Qy      401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db     1826 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1885

Qy      421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db     1886 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1945

Qy      441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db     1946 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2005

Qy      461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db     2006 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2065

Qy      481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db     2066 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2125

Qy      501 ValGlnAlaAlaLeuGluLys 507
Db     2126 GTTCAGGTGGCTCTGGATAAG 2146
```

Search completed: March 31, 2003, 16:15:25
Job time : 246.092 secs

GenCore version 5.1.4_p5_4578


```
OM protein - nucleic search, using frame_plus_p2n model

Run on:         March 30, 2003, 01:24:25 ; Search time 6948.72 Seconds
                                           (without alignments)
                                           2152.747 Million cell updates/sec

Title:          US-09-873-409-8
Perfect score:  2592
Sequence:       1 MILGILASLVNGACLPLMPL.........SESKSAVQAALEKDTPRYSF 514

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters:      4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115418_19228/app_query.fasta_1.7544
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_54887_@runat_27032003_115418_19228 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      GenEmbl:*
                1:  gb_ba:*
                2:  gb_htg:*
                3:  gb_in:*
                4:  gb_om:*
                5:  gb_ov:*
                6:  gb_pat:*
                7:  gb_ph:*
                8:  gb_pl:*
                9:  gb_pr:*
                10:  gb_ro:*
                11:  gb_sts:*
                12:  gb_sy:*
                13:  gb_un:*
                14:  gb_vi:*
                15:  em_ba:*
                16:  em_fun:*
                17:  em_hum:*
                18:  em_in:*
                19:  em_mu:*
                20:  em_om:*
                21:  em_or:*
                22:  em_ov:*
                23:  em_pat:*
                24:  em_ph:*
                25:  em_pl:*
                26:  em_ro:*
                27:  em_sts:*
                28:  em_un:*
                29:  em_vi:*
                30:  em_htg_hum:*
                31:  em_htg_inv:*
                32:  em_htg_other:*
                33:  em_htg_mus:*
                34:  em_htg_pln:*
                35:  em_htg_rod:*
                36:  em_htg_mam:*
                37:  em_htg_vrt:*
                38:  em_sy:*
                39:  em_htgo_hum:*
                40:  em_htgo_mus:*
                41:  em_htgo_other:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2592	100.0	1940	6	AX339034	AX339034 Sequence
2	2568.5	99.1	2021	6	AX339033	AX339033 Sequence
3	2552	98.5	3621	6	AX339032	AX339032 Sequence
4	2528.5	97.6	3702	6	AX339031	AX339031 Sequence
5	2460.5	94.9	3699	6	AX478104	AX478104 Sequence
6	1699.5	65.6	3177	6	AX339030	AX339030 Sequence
7	1590	61.3	4189	6	AX322791	AX322791 Sequence
8	1590	61.3	4298	10	MUSMDR	M14757 Mouse multi
9	1585	61.1	4279	6	AX105082	AX105082 Sequence
10	1581	61.0	4279	6	AX105078	AX105078 Sequence
11	1581	61.0	4279	6	AX105080	AX105080 Sequence
12	1580.5	61.0	4195	6	AX108656	AX108656 Sequence
13	1580	61.0	4045	12	AF269224	AF269224 Synthetic
14	1579	60.9	3860	6	AX322787	AX322787 Sequence
15	1579	60.9	4378	6	E02326	E02326 Multidrug r
16	1579	60.9	4646	6	AX391099	AX391099 Sequence
17	1578.5	60.9	4280	10	CRUPGPII	M60041 C.griseus P
18	1576	60.8	3888	10	AY082609	AY082609 Rattus no
19	1576	60.8	4233	6	AR123273	AR123273 Sequence
20	1576	60.8	4279	6	AX105057	AX105057 Sequence
21	1576	60.8	4317	4	AF045016	AF045016 Canis fam
22	1576	60.8	4317	6	AX105059	AX105059 Sequence
23	1576	60.8	4669	6	AR091275	AR091275 Sequence
24	1576	60.8	4669	6	AR203322	AR203322 Sequence
25	1572	60.6	4186	6	AX108654	AX108654 Sequence
26	1571.5	60.6	4254	6	AX401748	AX401748 Sequence
27	1571.5	60.6	4254	10	RATMDRM	M81855 Rat mdr mRN
28	1571.5	60.6	4927	10	AF257746	AF257746 Rattus no
29	1570.5	60.6	4323	10	AF286167	AF286167 Rattus no
30	1570	60.6	3860	6	AX322789	AX322789 Sequence
31	1570	60.6	4646	6	AX336420	AX336420 Sequence
32	1570	60.6	4646	6	AX336708	AX336708 Sequence
33	1570	60.6	4646	6	I49610	I49610 Sequence 2
34	1570	60.6	4646	9	HUMMDR1	M14758 Homo sapien
35	1570	60.6	4669	6	I08557	I08557 Sequence 3
36	1570	60.6	6505	6	AR028671	AR028671 Sequence
37	1570	60.6	8630	6	AX012320	AX012320 Sequence
38	1570	60.6	8630	6	AX012321	AX012321 Sequence
39	1570	60.6	9318	6	AR028672	AR028672 Sequence
40	1568.5	60.5	4264	6	AR051647	AR051647 Sequence
41	1568.5	60.5	4264	6	AR051650	AR051650 Sequence
42	1565.5	60.4	4356	10	MUSMDRAA	M30697 Mouse multi
43	1565	60.4	4296	10	CRUPGPI	M60040 C.griseus P
44	1564	60.3	4304	10	CRUPGP1165	M59253 Chinese ham
45	1563.5	60.3	3489	4	AB029153	AB029153 Felis cat

ALIGNMENTS

RESULT 1

```
AX339034
LOCUS       AX339034                1940 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 16 from Patent WO0194400.
ACCESSION   AX339034
VERSION     AX339034.1  GI:18129126
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 16 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..1940
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      530 a    394 c    448 g    568 t
ORIGIN

Alignment Scores:
Pred. No.:               2.68e-192      Length:         1940
Score:                   2592.00        Matches:        514
Percent Similarity:      100.00%        Conservative:   0
Best Local Similarity:   100.00%        Mismatches:     0
Query Match:             100.00%        Indels:         0
DB:                      6              Gaps:           0

US-09-873-409-8 (1-514) x AX339034 (1-1940)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     34 ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 93

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153

Qy     41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    154 TCTTTCTTCAGGTTGACCCTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGT 213

Qy     61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    214 TACATACAGATTTCCTTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAA 273

Qy     81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    274 CAGTTTTTTCATTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333

Qy    101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    334 GAACTTAACACTCGCATGACAGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATT 393

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAG 453

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    454 GGCTGGAAACTCACCCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA 513

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAG 633

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    634 GAGAAAGAACTTCAAAGGTATACACAGAATCTCAAAGATGCAAAGGATTTTGGCATAAAA 693

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    694 AGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAATGGAACCTAT 753

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    754 GGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGATATACCATC 813

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    814 GGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATTGGAGCAGCA 873

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    874 GTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTT 933

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    934 ATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATA 993

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    994 GAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAG 1053

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1054 ATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTC 1113

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1114 AATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGAT 1173

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1174 GGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGAC 1233

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1234 CATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATC 1293

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1294 AAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAAT 1353

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1354 GCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGA 1413

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1414 GCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAAC 1473

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1474 CCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCT 1533

Qy    501 ValGlnAlaAlaLeuGluLysAspThrProArgTyrSerPhe 514
          ||||||||||||||||||||||||||||||||||||||||||
Db   1534 GTTCAAGCTGCACTGGAGAAGGATACCCCCAGGTATTCATTT 1575

RESULT 2
AX339033
LOCUS       AX339033                2021 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 15 from Patent WO0194400.
ACCESSION   AX339033
VERSION     AX339033.1  GI:18129125
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
```

```
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1
  AUTHORS    Frank,M.H. and Sayegh,M.H.
  TITLE      A gene encoding a multidrug resistance human p-glycoprotein
             homologue on chromosome 7p15-21 and uses thereof
  JOURNAL    Patent: WO 0194400-A 15 13-DEC-2001;
             THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..2021
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT      547 a    406 c    462 g    605 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              1.89e-190     Length:         2021
Score:                  2568.50       Matches:        514
Percent Similarity:     95.01%        Conservative:   0
Best Local Similarity:  95.01%        Mismatches:     0
Query Match:            99.09%        Indels:         27
DB:                     6             Gaps:           1

US-09-873-409-8 (1-514) x AX339033 (1-2021)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     34 ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 93

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153

Qy     41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    154 TCTTTCTTCAGGTTGACCCTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGT 213

Qy     61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    214 TACATACAGATTTCCTTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAA 273

Qy     81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    274 CAGTTTTTTCATTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333

Qy    101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    334 GAACTTAACACTCGCATGACAGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATT 393

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAG 453

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    454 GGCTGGAAACTCACCCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA 513

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAG 633

Qy    201 GluLysGluLeuGln--------------------------------------------- 205
          |||||||||||||||
Db    634 GAGAAAGAACTTCAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTTGGTTTTATTTT 693

Qy    206 ------------------------------------ArgTyrThrGlnAsnLeuLysAsp 213
                                              ||||||||||||||||||||||||
Db    694 CCCCAGTGGCTACTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGAATCTCAAAGAT 753

Qy    214 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 233
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    754 GCAAAGGATTTTGGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTAC 813

Qy    234 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 253
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    814 TTCTTTATGAATGGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAAT 873

Qy    254 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 273
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    874 GGAGAACCTGGATATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGC 933

Qy    274 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 293
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    934 AGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCT 993

Qy    294 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 313
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    994 GCCTTTCATATTTTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCT 1053

Qy    314 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 333
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1054 GGATATAAACCTGAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTAT 1113

Qy    334 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 353
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1114 CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAG 1173

Qy    354 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 373
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1174 ACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG 1233

Qy    374 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 393
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1234 AGGTTATATGATCCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1293

Qy    394 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 413
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1294 AATGTGCGGCATTATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGG 1353

Qy    414 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 433
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1354 ACCACCATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1413

Qy    434 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 453
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1414 AGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAAT 1473

Qy    454 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 473
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1474 ACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATT 1533

Qy    474 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 493
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1534 GCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTG 1593

Qy    494 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLysAspThrProArgTyrSer 513
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1594 GATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTGGAGAAGGATACCCCCAGGTATTCA 1653

Qy    514 Phe 514
          |||
Db   1654 TTT 1656

RESULT 3
AX339032
LOCUS       AX339032                3621 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 14 from Patent WO0194400.
ACCESSION   AX339032
VERSION     AX339032.1  GI:18129124
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
```

```
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 14 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..3621
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1081 a    696 c    842 g   1002 t
ORIGIN

Alignment Scores:
Pred. No.:              7.4e-189       Length:          3621
Score:                  2552.00        Matches:         507
Percent Similarity:     100.00%        Conservative:    0
Best Local Similarity:  100.00%        Mismatches:      0
Query Match:            98.46%         Indels:          0
DB:                     6              Gaps:            0

US-09-873-409-8 (1-514) x AX339032 (1-3621)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       34 ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 93

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153

Qy       41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      154 TCTTTCTTCAGGTTGACCCTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGT 213

Qy       61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      214 TACATACAGATTTCCTTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAA 273

Qy       81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      274 CAGTTTTTTCATTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333

Qy      101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      334 GAACTTAACACTCGCATGACAGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATT 393

Qy      121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAG 453

Qy      141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      454 GGCTGGAAACTCACCCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA 513

Qy      161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573

Qy      181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAG 633

Qy      201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      634 GAGAAAGAACTTCAAAGGTATACACAGAATCTCAAAGATGCAAAGGATTTTGGCATAAAA 693

Qy      221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      694 AGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAATGGAACCTAT 753

Qy      241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      754 GGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGATATACCATC 813

Qy      261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      814 GGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATTGGAGCAGCA 873

Qy      281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      874 GTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTT 933

Qy      301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      934 ATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATA 993

Qy      321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      994 GAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAG 1053

Qy      341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1054 ATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTC 1113

Qy      361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1114 AATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGAT 1173

Qy      381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1174 GGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGAC 1233

Qy      401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1234 CATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATC 1293

Qy      421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1294 AAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAAT 1353

Qy      441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1354 GCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGA 1413

Qy      461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1414 GCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAAC 1473

Qy      481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1474 CCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCT 1533

Qy      501 ValGlnAlaAlaLeuGluLys 507
              |||||||||||||||||||||
Db     1534 GTTCAAGCTGCACTGGAGAAG 1554

RESULT 4
AX339031
LOCUS       AX339031             3702 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 13 from Patent WO0194400.
ACCESSION   AX339031
VERSION     AX339031.1  GI:18129123
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 13 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
```

```
FEATURES             Location/Qualifiers
     source          1. .3702
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1098 a    708 c    856 g   1039 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:                5.1e-187        Length:         3702
Score:                    2528.50         Matches:        507
Percent Similarity:       94.94%          Conservative:   0
Best Local Similarity:    94.94%          Mismatches:     0
Query Match:              97.55%          Indels:         27
DB:                       6               Gaps:           1

US-09-873-409-8 (1-514) x AX339031 (1-3702)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      34 ATGATCCTGGGTATACTGGCATCACTGGTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 93

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      94 GTTTTAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCAAACTAACACATAC 153

Qy      41 SerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGly 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     154 TCTTTCTTCAGGTTGACCCTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGT 213

Qy      61 TyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLys 80
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     214 TACATACAGATTTCCTTGTGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAA 273

Qy      81 GlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGly 100
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     274 CAGTTTTTTCATTCAGTTTTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGT 333

Qy     101 GluLeuAsnThrArgMetThrAspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     334 GAACTTAACACTCGCATGACAGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATT 393

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     394 GCTCTGTTGTTTCAAAACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAG 453

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     454 GGCTGGAAACTCACCCTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCA 513

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     514 GCATGTTCTAGGATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCT 573

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     574 GGGGCTGTGGCAGAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAG 633

Qy     201 GluLysGluLeuGln--------------------------------------------- 205
           |||||||||||||||
Db     634 GAGAAAGAACTTCAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTTGGTTTTATTTT 693

Qy     206 ------------------------------------ArgTyrThrGlnAsnLeuLysAsp 213
                                               ||||||||||||||||||||||||
Db     694 CCCCAGTGGCTACTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGAATCTCAAAGAT 753

Qy     214 AlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyr 233
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     754 GCAAAGGATTTTGGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTAC 813

Qy     234 PhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsn 253
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     814 TTCTTTATGAATGGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAAT 873

Qy     254 GlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSer 273
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     874 GGAGAACCTGGATATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGC 933

Qy     274 SerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAla 293
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     934 AGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCT 993

Qy     294 AlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAla 313
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     994 GCCTTTCATATTTTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCT 1053

Qy     314 GlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyr 333
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1054 GGATATAAACCTGAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTAT 1113

Qy     334 ProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGlu 353
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1114 CCATCAAGACCATCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAG 1173

Qy     354 ThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGln 373
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1174 ACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAG 1233

Qy     374 ArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeu 393
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1234 AGGTTATATGATCCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTA 1293

Qy     394 AsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGly 413
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1294 AATGTGCGGCATTATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGG 1353

Qy     414 ThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGlu 433
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1354 ACCACCATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAG 1413

Qy     434 ArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsn 453
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1414 AGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAAT 1473

Qy     454 ThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIle 473
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1474 ACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATT 1533

Qy     474 AlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeu 493
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1534 GCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTG 1593

Qy     494 AspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 507
           ||||||||||||||||||||||||||||||||||||||||||
Db    1594 GATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTGGAGAAG 1635

RESULT 5
AX478104
LOCUS       AX478104             3699 bp    DNA     linear   PAT 12-AUG-2002
DEFINITION  Sequence 34 from Patent WO0240541.
ACCESSION   AX478104
VERSION     AX478104.1  GI:22217064
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Tang,Y.T., Yue,H., Nguyen,D.B., Hafalia,A.J., Elliott,V.S., Lu,Y.,
            Walia,N.K., Yao,M.G., Baughn,M.R., Gandhi,A.R., Ding,L.,
            Sanjanwala,M., Ramkumar,J., Arvizu,C., Gietzen,K.J., Lal,P.G.,
            Azimzai,Y., Khan,F.A., Thangavelu,K., Thornton,M., Lu,D.A.,
            Tribouley,C.M., Warren,B.A., Ison,C.H., Das,D., Raumann,B.E.,
            Policky,J.L. and Kearney,L.
  TITLE     Transporters and ion channels
  JOURNAL   Patent: WO 0240541-A 34 23-MAY-2002;
            Incyte Genomics, Inc. (US)
```

```
FEATURES             Location/Qualifiers
     source          1..3699
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /note="Incyte ID No: 7472030CB1"
BASE COUNT     1116 a    707 c    860 g   1016 t
ORIGIN

Alignment Scores:
Pred. No.:              9.79e-182     Length:          3699
Score:                  2460.50       Matches:         495
Percent Similarity:     95.40%        Conservative:    3
Best Local Similarity:  94.83%        Mismatches:      9
Query Match:            94.93%        Indels:          15
DB:                     6             Gaps:            2

US-09-873-409-8 (1-514) x AX478104 (1-3699)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           ||||||||||||||||    ||||||    |||||||||||||||||||||||||||
Db     145 ATGATCCTGGGTATACTGACATCACTGTTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 204

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
              :::|||||||||||||||||||||||||||||||||||||   ||||||||
Db     205 TGTATAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCACACTAACACAACA 264

Qy      41 SerPhePheArg-----------------------------------------LeuThr 46
           ::::::                                                |||||
Db     265 AATTATCAGAACTGTACTCAGTCTCAAGAGAAGCTGAATGAAGATATGACTCTGTTGACC 324

Qy      47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     325 CTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGTTACATACAGATTTCCTTG 384

Qy      67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     385 TGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAACAGTTTTTCATTCAGTT 444

Qy      87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     445 TTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGTGAACTTAACACTCGCATG 504

Qy     107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
           |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     505 ACAGATGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATTGCTCTGTTGTTTCAA 564

Qy     126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     565 AACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAGGGCTGGAAACTCACC 624

Qy     146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     625 CTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCAGCATGTTCTAGGATG 684

Qy     166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     685 GTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCAGAA 744

Qy     186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     745 GAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAGGAGAAAGAACTTCAA 804

Qy     206 ArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSer 225
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     805 AGGTATACACAGAATCTCAAAGATGCAAAGGATTTTGGCATAAAAAGGACTATAGCTTCA 864

Qy     226 LysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrp 245
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     865 AAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAATGGAACCTATGGACTTGCTTTTTGG 924

Qy     246 TyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAla 265
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     925 TATGGAACCTCCTTGATTCTTAATGGAGAACCTGGATATACCATCGGGACTGTTCTTGCT 984

Qy     266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     985 GTTTTCTTTAGTGTAATCCATAGTAGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAA 1044

Qy     286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1045 ACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTTATTGATAAGAAACCC 1104

Qy     306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
           ||||||    ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1105 AGTATAGGTAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGAA 1164

Qy     326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1165 TTTAAAAATGTTTCTTTCAATTATCCATCAAGGCCATCTATCAAGATTCTGAAAGGTCTG 1224

Qy     346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
           ||||||    ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1225 AATCTCGGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAG 1284

Qy     366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1285 AGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGTG 1344

Qy     386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1345 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTGGTT 1404

Qy     406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1405 AGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGAGAT 1464

Qy     426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1465 GATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATC 1524

Qy     446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1525 ATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGA 1584

Qy     466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1585 GGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATT 1644

Qy     486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeu 505
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1645 TTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTG 1704

Qy     506 GluLys 507
           ||||||
Db    1705 GAGAAG 1710

RESULT 6
AX339030
LOCUS       AX339030                3177 bp    DNA     linear   PAT 09-JAN-2002
DEFINITION  Sequence 12 from Patent WO0194400.
ACCESSION   AX339030
VERSION     AX339030.1  GI:18129122
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Frank,M.H. and Sayegh,M.H.
  TITLE     A gene encoding a multidrug resistance human p-glycoprotein
            homologue on chromosome 7p15-21 and uses thereof
  JOURNAL   Patent: WO 0194400-A 12 13-DEC-2001;
            THE BRIGHAM AND WOMEN'S HOSPITAL, INC. (US)
FEATURES             Location/Qualifiers
     source          1..3177
                     /organism="Homo sapiens"
```

```
                    /db_xref="taxon:9606"
BASE COUNT      970 a    601 c     735 g     870 t       1 others
ORIGIN

Alignment Scores:
Pred. No.:                1.1e-122       Length:         3177
Score:                    1699.50        Matches:        343
Percent Similarity:       92.70%         Conservative:   0
Best Local Similarity:    92.70%         Mismatches:     0
Query Match:              65.57%         Indels:         27
DB:                       6              Gaps:           1

US-09-873-409-8 (1-514) x AX339030 (1-3177)

Qy      165 MetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAla 184
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        1 ATGGTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCA 60

Qy      185 GluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeu 204
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       61 GAAGAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAGGAGAAAGAACTT 120

Qy      205 Gln--------------------------------------------------------- 205
            |||
Db      121 CAAAGGTCTTTCCTTTTAAATATAACAAGATATGCTTGGTTTTATTTTCCCCAGTGGCTA 180

Qy      206 ------------------------ArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 217
                                    ||||||||||||||||||||||||||||||||||||
Db      181 CTAAGTTGTGTTCTGTTNTTTGTAAGGTATACACAGAATCTCAAAGATGCAAAGGATTTT 240

Qy      218 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 237
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      241 GGCATAAAAAGGACTATAGCTTCAAAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAAT 300

Qy      238 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 257
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      301 GGAACCTATGGACTTGCTTTTTGGTATGGAACCTCCTTGATTCTTAATGGAGAACCTGGA 360

Qy      258 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 277
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      361 TATACCATCGGGACTGTTCTTGCTGTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATT 420

Qy      278 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 297
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      421 GGAGCAGCAGTCCCTCACTTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATT 480

Qy      298 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 317
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      481 TTCCAGGTTATTGATAAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCT 540

Qy      318 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 337
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      541 GAATCCATAGAAGGAACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCA 600

Qy      338 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 357
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      601 TCTATCAAGATTCTGAAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTG 660

Qy      358 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 377
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      661 GTCGGTCTCAATGGCAGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGAT 720

Qy      378 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 397
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      721 CCGGATGATGGCTTTATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCAT 780

Qy      398 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 417
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      781 TATCGAGACCATATTGGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGT 840

Qy      418 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 437
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      841 AACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGG 900

Qy      438 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 457
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      901 GAAGCAAATGCGTATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGG 960

Qy      458 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 477
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      961 GAAAAAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTA 1020

Qy      478 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 497
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1021 GTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGC 1080

Qy      498 LysSerAlaValGlnAlaAlaLeuGluLys 507
            ||||||||||||||||||||||||||||||
Db     1081 AAGTCAGCTGTTCAAGCTGCACTGGAGAAG 1110

RESULT 7
AX322791
LOCUS       AX322791                 4189 bp    DNA     linear   PAT 07-JAN-2002
DEFINITION  Sequence 5 from Patent WO0192877.
ACCESSION   AX322791
VERSION     AX322791.1  GI:18093768
KEYWORDS    .
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Sorrentino,B. and Schuetz,J.
  TITLE     Method of identifying and/or isolating stem cells
  JOURNAL   Patent: WO 0192877-A 5 06-DEC-2001;
            ST. JUDE CHILDREN'S RESEARCH HOSPITAL (US)
FEATURES             Location/Qualifiers
     source          1. .4189
                     /organism="Mus musculus"
                     /db_xref="taxon:10090"
BASE COUNT     1204 a    875 c    1028 g    1082 t
ORIGIN

Alignment Scores:
Pred. No.:                4.99e-114      Length:         4189
Score:                    1590.00        Matches:        307
Percent Similarity:       76.32%         Conservative:   99
Best Local Similarity:    57.71%         Mismatches:     94
Query Match:              61.34%         Indels:         32
DB:                       6              Gaps:           5

US-09-873-409-8 (1-514) x AX322791 (1-4189)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            |||||||||||    |||||:::::::::::|||         ||||||||:::    |||
Db      148 ATGATTCTGGGAACTCTCGCTGCTATTATCCATGGAACATTACTTCCCCTCTTGATGCTG 207

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
            |||   |||   |||:::|||:::
Db      208 GTGTTTGGAAACATGACAGATAGTTTTACAAAAGCAGAAGCCAGTATTCTGCCAAGCATT 267

Qy       31 ------------------SerGlyCysLeuValGlnThrAsnThr--------------- 39
                              :::    |||:::   :::|||:::
Db      268 ACTAATCAAAGTGGACCCAACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAAGAG 327

Qy       40 ---------TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56
                     |||::::::               ||||||   |||||||||
Db      328 ATGGCCATATACGCCTAC---------------TATTACACCGGGATTGGTGCTGGTGTG 372

Qy       57 LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys 76
            ||||||   |||||||||:::|||||||||   :::   |||   ||||||
Db      373 CTCATAGTTGCCTACATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGAAGACAGATACAC 432

Qy       77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96
            :::||||||::::::||||||||||:::::::::   |||:::|||||||||||||||
Db      433 AAGATTAGGCAGAAGTTTTTCCATGCTATAATGAATCAGGAGATAGGCTGGTTTGATGTG 492
```

```
Qy       97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
               |||:::||||||||||||||||||:::|||   |||:::  ||||||:::||||||
Db      493 CATGATGTTGGGGAGCTCAACACCCGGCTCACAGATGATGTCTCCAAAATTAATGACGGA 552

Qy      116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
            ||||||||||||||    :::   ||||||:::::::::||||||      |||
Db      553 ATTGGTGACAAAATTGGGATGTTTTTCAGTCCATAACCACATTTTAGCCGGTTTTATC 612

Qy      136 ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
            :::|||    :::   ||||||||||||||||||||||    |||:::   |||||||||||
Db      613 ATAGGATTTATAAGTGGTTGGAAGCTAACCCTTGTCATTTTGGCTGTCAGCCCTCTTATT 672

Qy      156 MetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
                  |||:::|||     :::::::::::   |||    |||:::||||||||
Db      673 GGATTGTCATCTGCTTTGTGGGCAAAGGTATTGACTTCATTTACTAATAAGGAACTCCAG 732

Qy      176 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIle 195
            ||||||:::||||||||||||||||||||||||||||||:::::||||||||||||||||
Db      733 GCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT 792

Qy      196 AlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLys 215
            ||||||      |||:::|||||||||:::||||||   :::||||||::::::||||||
Db      793 GCCTTTGGAGGACAACAGAAGGAACTTGAAAGGTACAATAAAAATTTAGAAGAAGCTAAA 852

Qy      216 AspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePhe 235
            :::   |||||||||:::   |||   :::   :::|||:::|||      |||
Db      853 AATGTTGGCATAAAGAAAGCTATCACAGCCAGCATTTCGATAGGCATTGCCTACCTGTTG 912

Qy      236 MetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGlu 255
            :::      :::|||   |||||||||||||||||||||||||||:::|||:::   |||
Db      913 GTCTATGCATCATATGCACTGGCATTCTGGTATGGGACATCCTTGGTCCTCTCAAATGAA 972

Qy      256 ProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyr 275
                  |||:::||||||   ||||||   ||||||||||||::::::      ::::::
Db      973 ------TATTCTATTGGAGAAGTGCTTACTGTCTTCTTCTCTATTTGTTGGGGACTTTT 1026

Qy      276 CysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPhe 295
               ||||||      |||:::   |||   ||||||   |||||||||||||||||
Db     1027 AGTATTGGACACTTGGCCCCAAACATAGAAGCCTTTGCAAACGCACGAGGGGCAGCCTTT 1086

Qy      296 HisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyr 315
               ||||||::::::||||||   :::||||||||||||:::|||||||||   ||||||
Db     1087 GAAATCTTCAAGATAATTGATAACGAGCCAAGCATTGACAGCTTCTCAACAAAGGGCTAC 1146

Qy      316 LysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSer 335
            ||||||:::||||||   |||   :::|||||||||||||||   |||||||||||||||
Db     1147 AAACCAGACAGTATAATGGGAAACTTAGAGTTTAAAAATGTTCACTTCAACTACCCATCG 1206

Qy      336 ArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrVal 355
            |||      ::::::||||||||||||||||||||||||||||:::::||||||:::||||
Db     1207 AGAAGCGAAGTTCAGATCTTGAAGGGCCTCAATCTGAAGGTGAAGAGCGGACAGACGGTG 1266

Qy      356 AlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeu 375
            ||||||||||||      :::|||   ||||||||||||   |||||||:::|||||||||
Db     1267 GCCTTGGTTGGCAACAGTGGCTGTGGAAAAAGCACAACTGTCCAGCTGATGCAGAGGCTC 1326

Qy      376 TyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVal 395
            ||||||||   :::|||   :::   :::|||       |||||||||   :::||||||
Db     1327 TACGACCCCCTGGAGGGCGTGGTCAGTATCGACGGACAAGACATCAGAACCATCAATGTG 1386

Qy      396 ArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThr 415
            |||:::   |||:::   |||||||||||||||||||||||||||||||   ||||||
Db     1387 AGGTATCTGAGGGAGATCATTGGTGTGGTGAGTCAGGAACCTGTGCTGTTTGCCACCACG 1446

Qy      416 IleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAla 435
            |||:::   ||||||:::||||||||||:::||||||||   :::|||:::|||:::|||
Db     1447 ATCGCCGAGAACATTCGCTATGGCCGAGAAGATGTCACCATGGATGAGATTGAGAAAGCT 1506

Qy      436 AlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeu 455
            :::||||||||||||||||||||||||||||:::   |||::::::|||:::||||||
Db     1507 GTCAAGGAAGCCAATGCCTATGACTTCATCATGAAACTGCCCCACCAATTTGACACCCTG 1566

Qy      456 ValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArg 475
            ||||||||:::||||||||:::|||||||||||||||||||||||||||||||||||||
Db     1567 GTTGGTGAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGAATCGCCATTGCCCGG 1626

Qy      476 AlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSer 495
            ||||||||||||||||||||||||:::|||||||||||||||||||||||||||:::
Db     1627 GCCCTGGTCCGCAATCCCAAGATCCTTTTGTTGGACGAGGCCACCTCAGCCCTGGATACA 1686

Qy      496 GluSerLysSerAlaValGlnAlaAlaLeuGluLys 507
            ||||||::::::   |||||||||||||||:::|||
Db     1687 GAAAGTGAAGCTGTGGTGCAGGCCGCACTGGATAAG 1722
```

RESULT 8
MUSMDR

```
LOCUS       MUSMDR                  4298 bp    mRNA    linear   ROD 04-OCT-1995
DEFINITION  Mouse multidrug resistant protein mRNA, complete cds.
ACCESSION   M14757
VERSION     M14757.1  GI:199100
KEYWORDS    membrane-associated glycoprotein; multidrug resistance protein.
SOURCE      Mus musculus (strain BALB/c, sub_species domesticus) (clone: lambda
            DR11) cDNA to mRNA.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 4298)
  AUTHORS   Gros,P., Croop,J. and Housman,D.
  TITLE     Mammalian multidrug resistance gene: complete cDNA sequence
            indicates strong homology to bacterial transport proteins
  JOURNAL   Cell 47 (3), 371-380 (1986)
  MEDLINE   87028229
   PUBMED   3768958
COMMENT     The mdr gene family includes two or more related but distinct
            cellular genes.
            There are 3 possible start codons besides the one given in the
            Features table, at positions 161-163, 221-223 and 257-259.  If the
            latter two start codons are used as initiator the first hydrophobic
            domain could possibly serve as a signal sequence.
FEATURES             Location/Qualifiers
     source          1..4298
                     /organism="Mus musculus"
                     /strain="BALB/c"
                     /sub_species="domesticus"
                     /db_xref="taxon:10090"
                     /clone="lambda DR11"
                     /cell_type="pre-B cell"
     mRNA            <1..4298
                     /product="mdr mRNA"
     CDS             110..3940
                     /codon_start=1
                     /product="multidrug resistance protein"
                     /protein_id="AAA79005.1"
                     /db_xref="GI:387426"
                     /translation="MEFEENLKGRADKNFSKMGKKSKKEKKEKKPAVGVFGMFRYADW
                     LDKLCMILGTLAAIIHGTLLPLLMLVFGNMTDSFTKAEASILPSITNQSGPNSTLIIS
                     NSSLEEEMAIYAYYYTGIGAGVLIVAYIQVSLWCLAAGRQIHKIRQKFFHAIMNQEIG
                     WFDVHDVGELNTRLTDDVSKINDGIGDKIGMFFQSITTFLAGFIIGFISGWKLTLVIL
                     AVSPLIGLSSALWAKVLTSFTNKELQAYAKAGAVAEEVLAAIRTVIAFGGQQKELERY
                     NKNLEEAKNVGIKKAITASISIGIAYLLVYASYALAFWYGTSLVLSNEYSIGEVLTVF
                     FSILLGTFSIGHLAPNIEAFANARGAAFEIFKIIDNEPSIDSFSTKGYKPDSIMGNLE
                     FKNVHFNYPSRSEVQILKGLNLKVKSGQTVALVGNSGCGKSTTVQLMQRLYDPLEGVV
                     SIDGQDIRTINVRYLREIIGVVSQEPVLFATTIAENIRYGREDVTMDEIEKAVKEANA
                     YDFIMKLPHQFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTESEA
                     VVQAALDKAREGRTTIVIAHRLSTVRNADVIAGFDGGVIVEQGNHDELMREKGIYFKL
                     VMTQTRGNEIEPGNNAYGSQSDTDASELTSEESKSPLIRRSIYRSVHRKQDQERRLSM
                     KEAVDEDVPLVSFWRILNLNLSEWPYLLVGVLCAVINGCIQPVFAIVFSRIVGVFSRD
                     DDHETKRQNCNLFSLFFLVMGLISFVTYFFQGFTFGKAGEILTKRVRYMVFKSMLRQD
                     ISWFDDHKNSTGSLTTRLASDASSVKGAMGARLAVVTQNVANLGTGVILSLVYGWQLT
                     LLLVVIIPLIVLGGIIEMKLLSGQALKDKKQLEISGKIATEAIENFRTIVSLTREQKF
                     ETMYAQSLQVPYRNAMKKAHVFGITFSFTQAMMYFSYAACFRFGAYLVAQQLMTFENV
                     MLVFSAVVFGAMAAGNTSSFAPDYAKAKVSASHIIRIIEKTPEIDSYSTEGLKPTLLE
                     GNVKFNGVQFNYPTRPNIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERFYDPM
                     AGSVFLDGKEIKQLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRAVSHEEIVRA
                     AKEANIHQFIDSLPDKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEATSAL
```

```
                DTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVIENGKVKEHGTHQQLLAQK
                GIYFSMVQAGAKRS"
BASE COUNT     1226 a    903 c   1059 g   1110 t
ORIGIN

Alignment Scores:
Pred. No.:              5.15e-114     Length:         4298
Score:                  1590.00       Matches:        307
Percent Similarity:     76.32%        Conservative:   99
Best Local Similarity:  57.71%        Mismatches:     94
Query Match:            61.34%        Indels:         32
DB:                     10            Gaps:           5

US-09-873-409-8 (1-514) x MUSMDR (1-4298)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
             |||||||||||  ||||||:::::::::::|||      |||||||||:::    |||
Db     257 ATGATTCTGGGAACTCTCGCTGCTATTATCCATGGAACATTACTTCCCCTCTTGATGCTG 316

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
           |||   |||   |||:::|||:::
Db     317 GTGTTTGGAAACATGACAGATAGTTTTACAAAAGCAGAAGCCAGTATTCTGCCAAGCATT 376

Qy      31 ------------------SerGlyCysLeuValGlnThrAsnThr--------------- 39
                             :::      |||:::   :::|||:::
Db     377 ACTAATCAAAGTGGACCCAACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAAGAG 436

Qy      40 ---------TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56
                    |||::::::                 ||||||  |||||||||
Db     437 ATGGCCATATACGCCTAC---------------TATTACACCGGGATTGGTGCTGGTGTG 481

Qy      57 LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys 76
           ||||||      |||||||||:::|||||||||     :::    |||    ||||||
Db     482 CTCATAGTTGCCTACATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGAAGACAGATACAC 541

Qy      77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96
           :::||||||::::::|||||||||:::::::::      |||:::||||||||||||||
Db     542 AAGATTAGGCAGAAGTTTTTCCATGCTATAATGAATCAGGAGATAGGCTGGTTTGATGTG 601

Qy      97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
              |||:::||||||||||||||||||:::|||      |||:::   ||||||:::||||||
Db     602 CATGATGTTGGGGAGCTCAACACCCGGCTCACAGATGATGTCTCCAAAATTAATGACGGA 661

Qy     116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
           ||||||||||||||     :::   ||||||::::::::::||||||      |||
Db     662 ATTGGTGACAAAATTGGGATGTTTTTTCAGTCCATAACCACATTTTTAGCCGGTTTTATC 721

Qy     136 ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
           :::|||    :::   ||||||||||||||||||||||   |||:::   ||||||||||||
Db     722 ATAGGATTTATAAGTGGTTGGAAGCTAACCCTTGTCATTTTGGCTGTCAGCCCTCTTATT 781

Qy     156 MetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
                     |||:::|||     :::::::::::   |||   |||:::||||||||
Db     782 GGATTGTCATCTGCTTTGTGGGCAAAGGTATTGACTTCATTTACTAATAAGGAACTCCAG 841

Qy     176 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIle 195
           ||||||:::|||||||||||||||||||||||||||||||:::::::||||||||||||||
Db     842 GCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT 901

Qy     196 AlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLys 215
           ||||||         |||:::|||||||||:::|||      :::|||||||:::::||||||
Db     902 GCCTTTGGAGGACAACAGAAGGAACTTGAAAGGTACAATAAAAATTTAGAAGAAGCTAAA 961

Qy     216 AspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePhe 235
           :::   |||||||||:::   |||   :::   :::|||:::|||       |||
Db     962 AATGTTGGCATAAAGAAAGCTATCACAGCCAGCATTTCGATAGGCATTGCCTACCTGTTG 1021

Qy     236 MetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGlu 255
           :::      :::|||   |||||||||||||||||||||||||||:::|||:::   |||
Db    1022 GTCTATGCATCATATGCACTGGCATTCTGGTATGGGACATCCTTGGTCCTCTCAAATGAA 1081

Qy     256 ProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyr 275
                    |||:::||||||   ||||||   ||||||||||||::::::       ::::::
Db    1082 ------TATTCTATTGGAGAAGTGCTTACTGTCTTCTTCTCTATTTTGTTGGGGACTTTT 1135

Qy     276 CysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPhe 295
              ||||||         |||:::   |||   ||||||   ||||||||||||||||
Db    1136 AGTATTGGACACTTGGCCCCAAACATAGAAGCCTTTGCAAACGCACGAGGGGCAGCCTTT 1195

Qy     296 HisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyr 315
              ||||||::::::||||||   :::||||||||||||:::|||||||||     ||||||
Db    1196 GAAATCTTCAAGATAATTGATAACGAGCCAAGCATTGACAGCTTCTCAACAAAGGGCTAC 1255

Qy     316 LysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSer 335
           ||||||:::||||||   |||   :::|||||||||||||||  |||||||||||||||
Db    1256 AAACCAGACAGTATAATGGGAAACTTAGAGTTTAAAAATGTTCACTTCAACTACCCATCG 1315

Qy     336 ArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrVal 355
           |||        ::::::||||||||||||||||||||||::::::|||||||||:::||||||
Db    1316 AGAAGCGAAGTTCAGATCTTGAAGGGCCTCAATCTGAAGGTGAAGAGCGGACAGACGGTG 1375

Qy     356 AlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeu 375
           |||||||||||      :::|||   ||||||||||||   |||||||||:::||||||||
Db    1376 GCCTTGGTTGGCAACAGTGGCTGTGGAAAAAGCACAACTGTCCAGCTGATGCAGAGGCTC 1435

Qy     376 TyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVal 395
           |||||||||      :::|||   :::   :::|||      |||||||||     :::||||||
Db    1436 TACGACCCCCTGGAGGGCGTGGTCAGTATCGACGGACAAGACATCAGAACCATCAATGTG 1495

Qy     396 ArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThr 415
           |||:::   ||||||||||||||||||||||||||||||||||||||||||     ||||||
Db    1496 AGGTATCTGAGGGAGATCATTGGTGTGGTGAGTCAGGAACCTGTGCTGTTTGCCACCACG 1555

Qy     416 IleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAla 435
           |||:::   ||||||:::||||||||||:::||||||||||     :::|||:::|||:::|||
Db    1556 ATCGCCGAGAACATTCGCTATGGCCGAGAAGATGTCACCATGGATGAGATTGAGAAAGCT 1615

Qy     436 AlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeu 455
              :::|||||||||||||||||||||||||||||||:::      |||::::::|||:::||||||
Db    1616 GTCAAGGAAGCCAATGCCTATGACTTCATCATGAAACTGCCCCACCAATTTGACACCCTG 1675

Qy     456 ValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArg 475
           |||||||||:::|||||||||:::||||||||||||||||||||||||||||||||||||
Db    1676 GTTGGTGAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGAATCGCCATTGCCCGG 1735

Qy     476 AlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSer 495
           |||||||||||||||||||||||||:::|||||||||||||||||||||||||||||:::
Db    1736 GCCCTGGTCCGCAATCCCAAGATCCTTTTGTTGGACGAGGCCACCTCAGCCCTGGATACA 1795

Qy     496 GluSerLysSerAlaValGlnAlaAlaLeuGluLys 507
           ||||||:::::::   ||||||||||||:::|||
Db    1796 GAAAGTGAAGCTGTGGTGCAGGCCGCACTGGATAAG 1831
```

```
RESULT 9
AX105082
LOCUS       AX105082               4279 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 26 from Patent WO0123540.
ACCESSION   AX105082
VERSION     AX105082.1  GI:13921232
KEYWORDS    .
SOURCE      dog.
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 4279)
  AUTHORS   Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and
            Patten,C.J.
  TITLE     P-glycoproteins and uses thereof
  JOURNAL   Patent: WO 0123540-A 26 05-APR-2001;
            GENTEST CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..4279
                     /organism="Canis familiaris"
                     /db_xref="taxon:9615"
     CDS             17..3862
                     /note="unnamed protein product"
```

```
                     /codon_start=1
                     /protein_id="CAC37723.1"
                     /db_xref="GI:13921233"
                     /translation="MDPEGGRKGSAEKNFWKMGKKSKKKEKKEKKPTVSTFAMFRYSN
                     WLDRLYMLVGTMAAIIHGAALPLMMLVFGNMTDSFANAGISRNKTFPVIINESITNNT
                     QHFINHLEEEMTTYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQILKIRKQFFHAIMRQ
                     EIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSIATFFTGFIVGFTRGWKLTL
                     VILAISPVLGLSAAIWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFGGQKKEL
                     ERYNKNLEEAKGIGIKKAITANISIGAAFLLIYASYALAFWYGTSLVLSSEYTIGQVL
                     TVFFSVLIGAFSIGQASPSIEAFANARGAAYEIFKIIDNKPSIDSYSKSGHKPDNIKG
                     NLEFKNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQRLYDPTD
                     GMVCIDGQDIRTINVRHLREITGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKE
                     ANAYDFIMKLPNKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTE
                     SEAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELMKEKGIY
                     FKLVTMQTRGNEIELENATGESKSESDALEMSPKDSGSSLIKRRSTRRSIHAPQGQDR
                     KLGTKEDLNENVPPVSFWRILKLNSTEWPYFVVGIFCAIINGGLQPAFSIIFSRIIGI
                     FTRDEDPETKRQNSNMFSVLFLVLGIISFITFFLQGFTFGKAGEILTKRLRYMVFRSM
                     LRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAVITQNIANLGTGIIISLIYG
                     WQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRTVVSLTR
                     EQKFEYMYAQSLQVPYRNSLRKAHIFGVSFSITQAMMYFSYAGCFRFGAYLVANEFMN
                     FQDVLLVFSAIVFGAMAVGQVSSFAPDYAKAKVSAAHVIMIIEKSPLIDSYSPHGLKP
                     NTLEGNVTFNEVVFNYPTRPDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERF
                     YDPLAGSVLIDGKEIKHLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSHEE
                     IVQAAKEANIHHFIETLPEKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEA
                     TSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGKVKEHGTHQQL
                     LAQKGIYFSMVSVQAGAKR"
BASE COUNT     1296 a    833 c    1009 g    1141 t
ORIGIN

Alignment Scores:
Pred. No.:              1.25e-113      Length:         4279
Score:                  1585.00        Matches:        304
Percent Similarity:     75.80%         Conservative:   97
Best Local Similarity:  57.47%         Mismatches:     104
Query Match:            61.15%         Indels:         24
DB:                     6              Gaps:           4

US-09-873-409-8 (1-514) x AX105082 (1-4279)

Qy     1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db   167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy    21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
Db   227 GTTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy    35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
Db   287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy    45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db   347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy    60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db   407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy    80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db   467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy   100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db   527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy   119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db   587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy   139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db   647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy   159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db   707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy   179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db   767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy   199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
Db   827 GGACAAAAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAGGAATTGGG 886

Qy   219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238
Db   887 ATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCA 946

Qy   239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
Db   947 TCATATGCTCTGGCTTTCTGGTATGGACCTCCTTGGTCCTCTCCAGTGAA------TAT 1000

Qy   259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
Db  1001 ACTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGA 1060

Qy   279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
Db  1061 CAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTC 1120

Qy   299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
Db  1121 AAGATAATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGAT 1180

Qy   319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
Db  1181 AATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAA 1240

Qy   339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
Db  1241 GTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTT 1300

Qy   359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
Db  1301 GGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCC 1360

Qy   379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
Db  1361 ACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTT 1420

Qy   399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
Db  1421 CGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAA 1480

Qy   419 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
Db  1481 AACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAA 1540

Qy   439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
Db  1541 GCCAATGCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAG 1600

Qy   459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
Db  1601 AGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTT 1660

Qy   479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
Db  1661 CGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAA 1720

Qy   499 SerAlaValGlnAlaAlaLeuGluLys 507
Db  1721 GCAGTGGTTCAGGTGGCCCTGGATAAG 1747
```

```
RESULT 10
AX105078
LOCUS       AX105078                4279 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 22 from Patent WO0123540.
ACCESSION   AX105078
VERSION     AX105078.1  GI:13921228
KEYWORDS    .
SOURCE      dog.
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 4279)
  AUTHORS   Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and
            Patten,C.J.
  TITLE     P-glycoproteins and uses thereof
  JOURNAL   Patent: WO 0123540-A 22 05-APR-2001;
            GENTEST CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..4279
                     /organism="Canis familiaris"
                     /db_xref="taxon:9615"
     CDS             17..3862
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC37721.1"
                     /db_xref="GI:13921229"
                     /translation="MDPEGGRKGSAEKNFWKMGKKSKKNEKKEKKPTVSTFAMFRYSN
                     WLDRLYMLVGTMAAIIHGAALPLMMLVFGNMTDSFANAGISRNKTFPVIINESITNNT
                     QHFINHLEEEMTTYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQILKIRKQFFHAIMRQ
                     EIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSIATFFTGFIVGFTRGWKLTL
                     VILAISPVLGLSAAIWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFGGQKKEL
                     ERYNKNLEEAKGIGIKKAITANISIGAAFLLIYASYALAFWYGTSLVLSSEYSIGQVL
                     TVFFSVLIGAFSIGQASPSIEAFANARGAAYEIFKIIDNKPSIDSYSKSGHKPDNIKG
                     NLEFKNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQRLYDPTD
                     GMVCIDGQDIRTINVRHLREITGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKE
                     ANAYDFIMKLPNKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTE
                     SEAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELMKEKGIY
                     FKLVTMQTRGNEIELENATGESKSESDALEMSPKDSGSSLIKRRSTRRSIHAPQGQDR
                     KLGTKEDLNENVPPVSFWRILKLNSTEWPYFVVGIFCAIINGGLQPAFSIIFSRIIGI
                     FTRDEDPETKRQNSNMFSVLFLVLGIISFITFFLQGFTFGKAGEILTKRLRYMVFRSM
                     LRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAVITQNIANLGTGIIISLIYG
                     WQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRTVVSLTR
                     EQKFEYMYAQSLQVPYRNSLRKAHIFGVSFSITQAMMYFSYAGCFRFGAYLVANEFMN
                     FQDVLLVFSAIVFGAMAVGQVSSFAPDYAKAKVSAAHVIMIIEKSPLIDSYSPHGLKP
                     NTLEGNVTFNEVVFNYPTRPDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERF
                     YDPLAGSVLIDGKEIKHLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSHEE
                     IMQAAKEANIHHFIETLPEKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEA
                     TSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGKVKEHGTHQQL
                     LAQKGIYFSMVSVQAGAKR"
BASE COUNT     1295 a    833 c    1008 g    1143 t
ORIGIN

Alignment Scores:
Pred. No.:             2.56e-113     Length:          4279
Score:                 1581.00       Matches:         303
Percent Similarity:    75.80%        Conservative:    98
Best Local Similarity: 57.28%        Mismatches:      104
Query Match:           61.00%        Indels:          24
DB:                    6             Gaps:            4

US-09-873-409-8 (1-514) x AX105078 (1-4279)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||::::::|||   :::|||::::::::::::|||||   |||||||||||   |||
Db     167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
           |||   |||   |||:::|||:::      :::      :::
Db     227 GTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy      35 --------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                               :::      ||||||      |||
Db     287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy      45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                          ||||||   |||||||||            |||:::
Db     347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy      60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
              |||||||||:::|||   |||   :::   |||   ||||||      :::||||||
Db     407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy      80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
           ||||||||||||||||:::::::::   |||:::|||||||||||||||||      |||:::
Db     467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy     100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
           ||||||||||||||||||:::|||   |||:::   ||||||::::::||||||||||||
Db     527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy     119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
           ||||||   :::   ||||||:::::::::||||||         |||      ||||||
Db     587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy     139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
              :::||||||||||||||||||||||||   |||:::   ||||||::::::   |||
Db     647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy     159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
           ||||||      ::::::::::::   |||   |||   |||||||||   ||||||:::
Db     707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy     179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
           |||||||||||||||||||||||||||:::::::||||||||||||||||||||
Db     767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy     199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
              |||:::|||||||||:::||||||   :::||||||::::::||||||      |||
Db     827 GGACAAAAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAGGAATTGGG 886

Qy     219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238
           ||||||:::   |||   :::   :::|||:::||||||   :::      :::
Db     887 ATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCA 946

Qy     239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
           :::|||   ||||||||||||||||||||||||||||:::|||:::   |||      |||
Db     947 TCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TAT 1000

Qy     259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
           :::||||||   ||||||   ||||||||||||||||:::      ::::::   ||||||
Db    1001 TCTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGA 1060

Qy     279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
              |||   |||      |||   ||||||   ||||||||||||||||:::   ||||||
Db    1061 CAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTC 1120

Qy     299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
           ::::::||||||   ||||||||||||||||||::::::|||   :::|||:::||||||:::
Db    1121 AAGATAATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGAT 1180

Qy     319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
           :::|||:::|||   :::|||||||||||||||   |||:::||||||||||||
Db    1181 AATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAA 1240

Qy     339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
           :::||||||||||||||||||||||||:::::::::|||||:::|||||||||||||||
Db    1241 GTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTT 1300

Qy     359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
           |||   :::|||   ||||||||||||   ||||||||:::||||||||||||||||||
Db    1301 GGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCC 1360

Qy     379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
              ||||||   :::   :::|||      ||||||||||   :::||||||||||||
Db    1361 ACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTT 1420
```

```
Qy      399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
            |||:::      ||||||||||||||||||||||||||||||    |||||||||:::
Db     1421 CGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAA 1480

Qy      419 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
            ||||||:::||||||||::::::|||||    :::|||:::|||:::|||    :::|||
Db     1481 AACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAA 1540

Qy      439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
            ||||||||||||||||||||||||:::    ||||||||||||:::||||||||||||||
Db     1541 GCCAATGCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAG 1600

Qy      459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
            :::|||||||||||:::|||||||||||||||||||||||||||||||||||||||||||
Db     1601 AGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTT 1660

Qy      479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
            |||||||||||||||||:::||||||||||||||||||||||||:::||||||:::
Db     1661 CGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAA 1720

Qy      499 SerAlaValGlnAlaAlaLeuGluLys 507
            :::   ||||||   ||||||:::|||
Db     1721 GCAGTGGTTCAGGTGGCCCTGGATAAG 1747
```

RESULT 11
AX105080

```
LOCUS       AX105080                4279 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 24 from Patent WO0123540.
ACCESSION   AX105080
VERSION     AX105080.1  GI:13921230
KEYWORDS    .
SOURCE      dog.
  ORGANISM  Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE   1  (bases 1 to 4279)
  AUTHORS   Stocker,P.J., Steimel-Crespi,D.T., Crespi,C.L., Reif,T.C. and
            Patten,C.J.
  TITLE     P-glycoproteins and uses thereof
  JOURNAL   Patent: WO 0123540-A 24 05-APR-2001;
            GENTEST CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..4279
                     /organism="Canis familiaris"
                     /db_xref="taxon:9615"
     CDS             17..3862
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC37722.1"
                     /db_xref="GI:13921231"
                     /translation="MDPEGGRKGSAEKNFWKMGKKSKKKEKKEKKPTVSTFAMFRYSN
                     WLDRLYMLVGTMAAIIHGAALPLMMLVFGNMTDSFANAGISRNKTFPVIINESITNNT
                     QHFINHLEEEMTTYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQILKIRKQFFHAIMRQ
                     EIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSIATFFTGFIVGFTRGWKLTL
                     VILAISPVLGLSAAIWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFGGQKKEL
                     ERYNKNLEEAKGIGIKKAITANISIGAAFLLIYASYALAFWYGTSLVLSSEYSIGQVL
                     TVFFSVLIGAFSIGQASPSIEAFANARGAAYEIFKIIDNKPSIDSYSKSGHKPDNIKG
                     NLEFKNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQRLYDPTD
                     GMVCIDGQDIRTINVRHLREITGVVSQEPVLFATTIAENIRYGRENVTMDEIEKAVKE
                     ANAYDFIMKLPNKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDTE
                     SEAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELMKEKGIY
                     FKLVTMQTRGNEIELENATGESKSESDALEMSPKDSGSSLIKRRSTRRSIHAPQGQDR
                     KLGTKEDLNENVPPVSFWRILKLNSTEWPYFVVGIFCAIINGGLQPAFSIIFSRIIGI
                     FTRDEDPETKRQNSNMFSVLFLVLGIISFITFFLQGFTFGKAGEILTKRLRYMVFRSM
                     LRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAVITQNIANLGTGIIISLIYG
                     WQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRTVVSLTR
                     EQKFEYMYAQSLQVPYRNSLRKAHIFGVSFSITQAMMYFSYAGCFRFGAYLVANEFMN
                     FQDVLLVFSAIVFGAMAVGQVSSFAPDYAKAKVSAAHVIMIIEKSPLIDSYSPHGLKP
                     NTLEGNVTFNEVVFNYPTRPDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLERF
                     YDPLAGSVLIDGKEIKHLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSRVVSHEE
                     IMQAAKEANIHHFIETLPEKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDEA
                     TSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGKVKEHGTHQQL
                     LAQKGIYFSMVSVQAGAKR"
BASE COUNT     1296 a    833 c   1008 g   1142 t
ORIGIN

Alignment Scores:
Pred. No.:              2.56e-113      Length:          4279
Score:                  1581.00        Matches:         303
Percent Similarity:     75.80%         Conservative:    98
Best Local Similarity:  57.28%         Mismatches:      104
Query Match:            61.00%         Indels:          24
DB:                     6              Gaps:            4

US-09-873-409-8 (1-514) x AX105080 (1-4279)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            |||::::::|||   :::|||::::::::::::||||||   |||||||||||   |||
Db      167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
            |||   |||   |||:::|||:::      :::      :::
Db      227 GTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy       35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                                 :::      ||||||      |||
Db      287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy       45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                           ||||||   ||||||||||         |||:::
Db      347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy       60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
               |||||||||:::|||   |||   :::   |||   |||||      :::||||||
Db      407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy       80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
            ||||||||||||||||:::::::::   |||:::||||||||||||||      |||:::
Db      467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy      100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
            ||||||||||||||||||:::|||   |||:::   ||||||::::::||||||||||||
Db      527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy      119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
            ||||||   :::   ||||||:::::::||||||      |||      ||||||
Db      587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy      139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
               :::|||||||||||||||||||   |||:::      ||||||::::::      |||
Db      647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy      159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
            ||||||      ::::::::::::   |||   |||   ||||||||||   ||||||:::
Db      707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy      179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
            ||||||||||||||||||||||||||||:::::::||||||||||||||||||||||||
Db      767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy      199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
               |||:::|||||||||:::||||||   :::||||||::::::||||||      |||
Db      827 GGACAAAAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAGGAATTGGG 886

Qy      219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238
            ||||||:::   |||   :::   :::|||:::||||||   :::      :::
Db      887 ATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCA 946

Qy      239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
            :::|||   ||||||||||||||||||||||||:::|||:::      |||      |||
Db      947 TCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TAT 1000

Qy      259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
            :::||||||   ||||||   |||||||||||||||:::         ::::::   ||||||
Db     1001 TCTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGA 1060
```

```
Qy     279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
              |||   |||      |||   |||||   ||||||||||||:::   |||||
Db    1061 CAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTC 1120

Qy     299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
           ::::::|||||   |||||||||||||:::::::|||    :::|||:::||||||:::
Db    1121 AAGATAATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGAT 1180

Qy     319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
           :::|||:::|||   :::||||||||||||||   |||:::||||||||||
Db    1181 AATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAA 1240

Qy     339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
           :::||||||||||||||||||||||||:::::::::||||||:::|||||||||||||||
Db    1241 GTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTT 1300

Qy     359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
           |||   :::|||   ||||||||||||   |||||||||:::||||||||||||||||||
Db    1301 GGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCC 1360

Qy     379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
              ||||||   :::   :::|||      |||||||||   :::||||||||||||
Db    1361 ACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTT 1420

Qy     399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
           |||:::      ||||||||||||||||||||||||||||||   |||||||||:::
Db    1421 CGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAA 1480

Qy     419 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
           ||||||:::||||||||||::::::||||||   :::|||:::|||:::|||   :::|||
Db    1481 AACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAA 1540

Qy     439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
           ||||||||||||||||||||||||:::   ||||||||||||:::|||||||||||||||
Db    1541 GCCAATGCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAG 1600

Qy     459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
           :::|||||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db    1601 AGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTT 1660

Qy     479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
           ||||||||||||||||||:::||||||||||||||||||||||||:::||||||:::
Db    1661 CGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAA 1720

Qy     499 SerAlaValGlnAlaAlaLeuGluLys 507
           :::   ||||||   ||||||:::|||
Db    1721 GCAGTGGTTCAGGTGGCCCTGGATAAG 1747

RESULT 12
AX108656
LOCUS       AX108656                4195 bp    DNA     linear   PAT 30-APR-2001
DEFINITION  Sequence 3 from Patent WO0123565.
ACCESSION   AX108656
VERSION     AX108656.1  GI:13923888
KEYWORDS    .
SOURCE      crab-eating macaque.
  ORGANISM  Macaca fascicularis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
            Cercopithecinae; Macaca.
REFERENCE   1  (bases 1 to 4195)
  AUTHORS   Stocker,P.J., Steimel-Crespi,D.T. and Crespi,C.L.
  TITLE     P-glycoproteins from macaca fascicularis and uses thereof
  JOURNAL   Patent: WO 0123565-A 3 05-APR-2001;
            GENTEST CORPORATION (US)
FEATURES             Location/Qualifiers
     source          1..4195
                     /organism="Macaca fascicularis"
                     /db_xref="taxon:9541"
     CDS             100..3951
                     /note="unnamed protein product"
                     /codon_start=1
                     /protein_id="CAC37765.1"
                     /db_xref="GI:13923889"
                     /translation="MDLEGDRNGGAEKKNFFKLNNKSKKDKKERKPTVSVFSMFRYSN
                     WLDKLYMVVGTLAAIIHGAGLPLMMLVFGDMTDTFANAGNLGDLGALLFNNTNSSNIT
                     DTVPVMNLEEDMTRYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQIHKIRKQFFHAIMR
                     QEIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSMATFFTGFIVGFTRGWKLT
                     LVILAISPVLGLSAAVWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFGGQKKE
                     LERYNKNLEEAKRIGIKKAITANISIGAAFLLIYASYALAFWYGTTLVLSKEYSIGQV
                     LTVFFSVLIGAFSVGQASPSIEAFANARGAAFEIFKIIDNKPSIDSYSKSGHKPDNIK
                     GNLEFRNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQRLYDPT
                     EGMVSVDGQDIRTINVRFLREIIGVVSQEPVLFATTIAENIRYGREDVTMDEIEKAVK
                     EANAYDFIMKLPQKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEATSALDT
                     ESEAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELMKEKGI
                     YFKLVTMQTAGNEIELENAADESKSEIDTLEMSSHDSGSSLIRKRSTRRSVRGSQGQD
                     RKLSTKEALDESIPPVSFWRIMKLNLTEWPYFVVGVFCAIINGGLQPAFAVIFSKIIG
                     IFTRNDDAETKRQNSNLFSLLFLVLGIVSFITFFLQGFTFGKAGEILTKRLRYMVFRS
                     MLRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAIITQNIANLGTGIIISLIY
                     GWQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRTVVSLT
                     QEQKFEHMYDQSLQVPYRNSLRKAHIFGITFSFTQAMMYFSYAGCFRFGAYLVAHSLM
                     SFEDVLLVFSAVVFGAMAVGQVSSFAPDYAKAKVSAAHIIMIIEKTPLIDSYSTEGLK
                     PNTLEGNVTFNEVVFNYPTRLDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVVQLLER
                     FYDPLAGKVLLDGKEIKQLNVQWLRAHLGIVSQEPILFDCSISENIAYGDNSRVVSQE
                     EIVRAAKEANIHAFIESLPNKYSTRVGDKGTQLSGGQKQRIAIARALVRQPHILLLDE
                     ATSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGRVKEHGTHQQ
                     LLAQKGIYFSMVSVQAGAKRQ"
BASE COUNT     1230 a    801 c    1039 g    1125 t
ORIGIN

Alignment Scores:
Pred. No.:              2.74e-113      Length:        4195
Score:                  1580.50        Matches:       306
Percent Similarity:     75.85%         Conservative:  96
Best Local Similarity:  57.74%         Mismatches:    103
Query Match:            60.98%         Indels:        25
DB:                     6              Gaps:          4

US-09-873-409-8 (1-514) x AX108656 (1-4195)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||::::::|||   ||||||::::::::::::||||||   ||||||||||||   |||
Db     250 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGAGCTGGACTTCCTCTCATGATGCTG 309

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSer---------------------GlyCys 33
           |||   |||:::|||:::|||         :::                     |||
Db     310 GTGTTTGGAGACATGACGGATACCTTTGCAAATGCAGGAAATTTAGGAGATTTAGGAGCT 369

Qy      34 LeuValGlnThrAsnThrTyrSer------------------------------------ 41
           |||:::      ||||||   |||
Db     370 CTGTTGTTTAACAACACTAATAGCAGTAATATCACTGATACAGTGCCCGTCATGAATCTG 429

Qy      42 ---------PhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIle 58
                    |||         ||||||   |||||||||         |||:::
Db     430 GAGGAAGATATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTT 489

Qy      59 PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIle 78
                 |||||||||:::|||   |||   :::   |||   ||||||      :::|||
Db     490 GCTGCTTACATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATT 549

Qy      79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp 98
           |||||||||||||||||||:::::::::   |||:::||||||||||||||||
Db     550 AGAAAACAGTTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGAT 609

Qy      99 IleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly 117
           :::||||||||||||||||||:::|||   |||:::   ||||||::::::||||||
Db     610 GTTGGGGAGCTTAACACCCGGCTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGT 669

Qy     118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
           |||||||||   :::   ||||||:::|||:::||||||         |||      ||||||
Db     670 GACAAAATTGGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGA 729

Qy     138 LeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
                 :::||||||||||||||||||   |||:::   ||||||::::::
Db     730 TTTACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTG 789
```

```
Qy      158 SerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr 177
Db      790 TCAGCTGCAGTCTGGGCAAAGATACTGTCTTCATTTACTGATAAAGAACTCTTAGCTTAT 849

Qy      178 SerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197
Db      850 GCAAAAGCTGGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTT 909

Qy      198 ArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 217
Db      910 GGAGGACAAAAGAAAGAACTCGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATT 969

Qy      218 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 237
Db      970 GGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTTATCTAT 1029

Qy      238 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 257
Db     1030 GCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAAAGGAA------ 1083

Qy      258 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 277
Db     1084 TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTT 1143

Qy      278 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 297
Db     1144 GGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTTTGAAATC 1203

Qy      298 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 317
Db     1204 TTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCA 1263

Qy      318 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 337
Db     1264 GATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAA 1323

Qy      338 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 357
Db     1324 GAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTG 1383

Qy      358 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 377
Db     1384 GTTGGAAACAGCGGCTGTGGGAAGAGCACAACGGTCCAGCTGATGCAGAGGCTTTATGAC 1443

Qy      378 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 397
Db     1444 CCCACAGAGGGCATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAACGTAAGGTTT 1503

Qy      398 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 417
Db     1504 CTACGGGAAATCATCGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCT 1563

Qy      418 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 437
Db     1564 GAAAACATTCGCTATGGTCGTGAAGATGTCACCATGGATGAGATTGAGAAAGCTGTCAAG 1623

Qy      438 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 457
Db     1624 GAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCAGAAATTTGACACCCTGGTTGGA 1683

Qy      458 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 477
Db     1684 GAGAGAGGGGCCCAGCTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTG 1743

Qy      478 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 497
Db     1744 GTTCGCAACCCCAAGATCCTCCTGCTGGACGAGGCCACGTCAGCCTTGGACACAGAAAGT 1803

Qy      498 LysSerAlaValGlnAlaAlaLeuGluLys 507
Db     1804 GAAGCAGTGGTTCAGGTGGCTCTGGATAAG 1833
```

```
RESULT 13
AF269224
LOCUS       AF269224                4045 bp    DNA     linear   SYN 04-JUL-2000
DEFINITION  Synthetic construct Canis familiaris his-tagged-multidrug
            resistance glycoprotein gene, complete cds.
ACCESSION   AF269224
VERSION     AF269224.1  GI:8926216
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 4045)
  AUTHORS   Roulet,A.C., Puel,O., Gesta,S., Alvinerie,M. and Pineau,T.J.
  TITLE     Analysis of dog MDR1 p-glycoprotein
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 4045)
  AUTHORS   Roulet,A.C., Puel,O., Gesta,S., Alvinerie,M. and Pineau,T.J.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-MAY-2000) Pharmacology, INRA, 180 Chemin de
            Tournefeuille, BP3, Toulouse 31931, France
FEATURES             Location/Qualifiers
     source          1..4045
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="original mRNA isolated from Beagle dog colon"
     CDS             101..3964
                     /codon_start=1
                     /transl_table=11
                     /product="his-tagged-multidrug resistance glycoprotein
                     MDR1"
                     /protein_id="AAF81747.1"
                     /db_xref="GI:8926217"
                     /translation="MHHHHHHDPEGGRKGSAEKNFWKMGKKSKKNEKKEKKPTVSTFA
                     MFRYSNWLDRLYMLVGTMAAIIHGAALPLMMLVFGNMTDSFANAGISRNKTFPVIINE
                     SITNNTQHFINHLEEEMTTYAYYYSGIGAGVLVAAYIQVSFWCLAAGRQILKIRKQFF
                     HAIMRQEIGWFDVHDVGELNTRLTDDVSKINEGIGDKIGMFFQSIATFFTGFIVGFTR
                     GWKLTLVILAISPVLGLSAAIWAKILSSFTDKELLAYAKAGAVAEEVLAAIRTVIAFG
                     GQKKELERYNKNLEEAKRIGIKKAITANISIGAAFLLIYASYALAFWYGTSLVLSSEY
                     SIGQVLTVFFSVLIGAFSIGQASPSIEAFANARGAAYEIFKIIDNKPSIDSYSKSGHK
                     PDNIKGNLEFKNVHFSYPSRKEVKILKGLNLKVQSGQTVALVGNSGCGKSTTVQLMQR
                     LYDPTDGMVCIDGQDIRTINVRHLREITGVVSQEPVLFATTIAENIRYGRENVTMDEI
                     EKAVKEANAYDFIMKLPNKFDTLVGERGAQLSGGQKQRIAIARALVRNPKILLLDEAT
                     SALDTESEAVVQVALDKARKGRTTIVIAHRLSTVRNADVIAGFDDGVIVEKGNHDELM
                     KEKGIYFKLVTMQTRGNEIELENATGESKSESDALEMSPKDSGSSLIKRRSTRRSIHA
                     PQGQDRKLGTKEDLNENVPPVSFWRILKLNSTEWPYFVVGIFCAIINGGLQPAFSIIF
                     SRIIGIFTRDEDPETKRQNSNMFSVLFLVLGIISFITFFLQGFTFGKAGEILTKRLRY
                     MVFRSMLRQDVSWFDDPKNTTGALTTRLANDAAQVKGAIGSRLAVITQNIANLGTGII
                     ISLIYGWQLTLLLLAIVPIIAIAGVVEMKMLSGQALKDKKELEGAGKIATEAIENFRT
                     VVSLTREQKFEYMYAQSLQVPYRNSLRKAHIFGVSFSITQAMMYFSYAGCFRFGAYLV
                     ANEFMNFQDVLLVFSAIVFGAMAVGQVSSFAPDYAKAKVSAAHVIMIIEKSPLIDSYS
                     PHGLKPNTLEGNVTFNEVVFNYPTRPDIPVLQGLSLEVKKGQTLALVGSSGCGKSTVV
                     QLLERFYDPLAGSVLIDGKEIKHLNVQWLRAHLGIVSQEPILFDCSIAENIAYGDNSR
                     VVSHEEIMQAAKEANIHHFIETLPEKYNTRVGDKGTQLSGGQKQRIAIARALVRQPHI
                     LLLDEATSALDTESEKVVQEALDKAREGRTCIVIAHRLSTIQNADLIVVFQNGKVKEH
                     GTHQQLLAQKGIYFSMISVQAGAKR"
BASE COUNT     1171 a    850 c    986 g   1038 t
ORIGIN

Alignment Scores:
Pred. No.:              2.86e-113      Length:         4045
Score:                  1580.00        Matches:        303
Percent Similarity:     75.80%         Conservative:   98
Best Local Similarity:  57.28%         Mismatches:     104
Query Match:            60.96%         Indels:         24
DB:                     12             Gaps:           4

US-09-873-409-8 (1-514) x AF269224 (1-4045)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db      269 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 328

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
```

```
Db     329 GTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 388

Qy      35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44

Db     389 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 448

Qy      45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59

Db     449 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 508

Qy      60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79

Db     509 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 568

Qy      80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99

Db     569 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 628

Qy     100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118

Db     629 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 688

Qy     119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138

Db     689 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 748

Qy     139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158

Db     749 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 808

Qy     159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178

Db     809 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 868

Qy     179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198

Db     869 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 928

Qy     199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218

Db     929 GGACAAAAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGG 988

Qy     219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238

Db     989 ATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCA 1048

Qy     239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258

Db    1049 TCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TAT 1102

Qy     259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278

Db    1103 TCTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGA 1162

Qy     279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298

Db    1163 CAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTC 1222

Qy     299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318

Db    1223 AAGATAATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGAT 1282

Qy     319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338

Db    1283 AATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAA 1342

Qy     339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358

Db    1343 GTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTT 1402

Qy     359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378

Db    1403 GGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCC 1462

Qy     379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398

Db    1463 ACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTT 1522

Qy     399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418

Db    1523 CGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAA 1582

Qy     419 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438

Db    1583 AACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAA 1642

Qy     439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458

Db    1643 GCCAATGCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAG 1702

Qy     459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478

Db    1703 AGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTT 1762

Qy     479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498

Db    1763 CGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAA 1822

Qy     499 SerAlaValGlnAlaAlaLeuGluLys 507

Db    1823 GCAGTGGTTCAGGTGGCCCTGGATAAG 1849

RESULT 14
AX322787
LOCUS       AX322787                3860 bp    DNA     linear   PAT 07-JAN-2002
DEFINITION  Sequence 1 from Patent WO0192877.
ACCESSION   AX322787
VERSION     AX322787.1  GI:18093766
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1
  AUTHORS   Sorrentino,B. and Schuetz,J.
  TITLE     Method of identifying and/or isolating stem cells
  JOURNAL   Patent: WO 0192877-A 1 06-DEC-2001;
            ST. JUDE CHILDREN'S RESEARCH HOSPITAL (US)
FEATURES             Location/Qualifiers
     source          1..3860
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1135 a    746 c    957 g   1022 t
ORIGIN

Alignment Scores:
Pred. No.:              3.23e-113     Length:         3860
Score:                  1579.00       Matches:        306
Percent Similarity:     77.04%        Conservative:   100
Best Local Similarity:  58.06%        Mismatches:     99
Query Match:            60.92%        Indels:         22
DB:                     6             Gaps:           5

US-09-873-409-8 (1-514) x AX322787 (1-3860)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20

Db     151 ATGGTGGTGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210

Qy      21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31

Db     211 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 270

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
```

```
Db    271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                |||          ||||||  ||||||||           |||:::    |||
Db    331 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 390

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
          ||||||:::|||   |||   :::   |||  ||||||      :::|||||||||||
Db    391 ATTCAGGTTTCATTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
          |||||||||:::::::::    |||:::||||||||||||||||      |||:::|||||
Db    451 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 510

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
          ||||||||||||:::|||    |||:::   ||||||::::::|||||||||||||||||
Db    511 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGGTATTGGTGACAAAATT 570

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
             :::  ||||||:::|||:::||||||       |||       ||||||      :::
Db    571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 630

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
          ||||||||||||||||||||||    |||:::   ||||||::::::      |||||||||
Db    631 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
             ::::::::::::    |||    |||    |||||||||    ||||||:::|||||||
Db    691 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
          ||||||||||||||||||||||||||:::::|||||||||||||||||||||      |||
Db    751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
          :::||||||||:::||||||    :::||||||::::::||||||       |||||||||
Db    811 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 870

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
          :::    |||    :::    :::|||:::||||||    :::       :::      :::|||
Db    871 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 930

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
             |||||||||||||||||||||:::|||:::|||:::||||||      |||:::|||
Db    931 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 984

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
          |||    ||||||    |||||||||||||||:::       ::::::    :::|||    |||
Db    985 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTAGTGTTGGACAGGCA 1044

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
             |||       |||    ||||||    ||||||||||||||||:::    ||||||::::::
Db   1045 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
          ||||||    |||||||||||||||:::::::|||    :::|||:::||||||::::::|||
Db   1105 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1164

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
          :::|||    :::|||||||:::||||||    |||:::||||||||||||      :::|||
Db   1165 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1224

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
          |||||||||||||||||||||||::::::::||||||:::||||||||||||||||||
Db   1225 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1284

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
          :::|||    |||||||||||    ||||||||:::||||||||||||||||||    :::
Db   1285 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1344

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
          |||    :::    ||||||       ||||||||    :::||||||||       |||:::
Db   1345 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1404

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
          ||||||||||||||||||||||||||||||    |||||||||:::    ||||||
Db   1405 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1464

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
          :::||||||||::::::||||||    :::|||:::|||:::|||    :::|||||||||
Db   1465 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1524

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
          ||||||||||||||||:::    |||:::||||||:::||||||||||||||:::|||
Db   1525 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1584

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
          ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1585 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1644

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
          ||||||||||:::||||||||||||||||||||||||||||:::||||||::::::
Db   1645 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1704

Qy    501 ValGlnAlaAlaLeuGluLys 507
          ||||||    ||||||:::|||
Db   1705 GTTCAGGTGGCTCTGGATAAG 1725
```

```
RESULT 15
E02326
LOCUS       E02326                  4378 bp    RNA     linear   PAT 29-SEP-1997
DEFINITION  Multidrug resistance relating gene derived from human normal cells.
ACCESSION   E02326
VERSION     E02326.1  GI:2170561
KEYWORDS    JP 1990100680-A/1.
SOURCE      Homo sapiens.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 4378)
  AUTHORS   Ueda,K. and Komano,T.
  TITLE     HUMAN NORMAL CELL-DERIVED MDR RELATED GENE
  JOURNAL   Patent: JP 1990100680-A 1 12-APR-1990;
            SUNTORY LTD
COMMENT     OS   Homo sapiens
            PN   JP 1990100680-A/1
            PD   12-APR-1990
            PF   05-OCT-1988 JP 1988251475
            PI   UEDA KAZUMITSU, KOMANO TORU
            PC   C12N15/12,C12N1/21,C12Q1/68;
            CC   strandedness: Single;
            CC   topology: Linear;
            CC   *source: tissue_type=Adrenal gland;
            CC   *source: clone=SAM1132;
            FH   Key             Location/Qualifiers
            FH
            FT   5'UTR           1..137
            FT   CDS             138..3980
            FT                   /gene='Multidrug resistance relating gene' FT
             mat_peptide     138..3977
            FT                   /gene='Multidrug resistance relating gene' FT
                             /gene='MDRA1'
            FT   3'UTR           3981..>4378.
FEATURES             Location/Qualifiers
     source          1..4378
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT     1315 a    818 c   1057 g   1188 t
ORIGIN

Alignment Scores:
Pred. No.:              3.77e-113      Length:         4378
Score:                  1579.00        Matches:        306
Percent Similarity:     77.04%         Conservative:   100
Best Local Similarity:  58.06%         Mismatches:     99
Query Match:            60.92%         Indels:         22
```

```
DB:                     6          Gaps:        5

US-09-873-409-8 (1-514) x E02326 (1-4378)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
             |||::::::|||   ||||||:::::::::::||||||   ||||||||||||    |||
Db     288 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 347

Qy      21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
             |||    ||||||||:::|||                         :::|||:::|||
Db     348 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 407

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                       ::::::       |||       ||||||
Db     408 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 467

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                  |||       ||||||   |||||||||         |||:::         |||
Db     468 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 527

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
             ||||||:::|||   |||   :::   |||   ||||||         :::||||||||||||
Db     528 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 587

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
             |||||||||:::::::::   |||:::||||||||||||||||||     |||:::||||||
Db     588 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 647

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
             ||||||||||||:::|||   |||:::   ||||||::::::|||||||||||||||||||
Db     648 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAAATT 707

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
                :::   ||||||:::|||:::||||||      |||      ||||||      :::
Db     708 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 767

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
             ||||||||||||||||||||   |||:::   ||||||::::::      |||||||||
Db     768 GGTTGGAAGCTAACCCTTGTGATTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 827

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
                   ::::::::::::   |||   |||   |||||||||   ||||||:::||||||
Db     828 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 887

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
             ||||||||||||||||||||||||:::::::||||||||||||||||||||      |||
Db     888 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 947

Qy     201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
             :::||||||||:::||||||   :::||||||::::::||||||      |||||||||
Db     948 AAGAAAGAACTTGAAAGGTACAACAAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1007

Qy     221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
             :::   |||   :::   :::|||:::||||||   :::      :::   :::|||
Db    1008 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1067

Qy     241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
                |||||||||||||||||||:::|||:::|||:::||||||      |||:::|||
Db    1068 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1121

Qy     261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
             |||   ||||||   |||||||||||||||:::      ::::::   :::|||   |||
Db    1122 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTAGTGTTGGACAGGCA 1181

Qy     281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
                |||      |||   ||||||   ||||||||||||||:::   ||||||::::::
Db    1182 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1241

Qy     301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
             ||||||   ||||||||||||||||:::::::|||   :::|||:::||||||::::::|||
Db    1242 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1301

Qy     321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
             :::|||   :::||||||:::||||||   |||:::||||||||||||      :::|||
Db    1302 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1361

Qy     341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
             ||||||||||||||||||||||:::::::::||||||:::|||||||||||||||||
Db    1362 ATCTTGAAGGGTCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1421

Qy     361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
             :::|||   ||||||||||||   |||||||||:::||||||||||||||||||   :::
Db    1422 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1481

Qy     381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
             |||   :::   ||||||      |||||||||   :::|||||||||      |||:::
Db    1482 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1541

Qy     401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
                |||||||||||||||||||||||||||||||||   |||||||||:::   ||||||
Db    1542 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1601

Qy     421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
             :::||||||||||::::::||||||   :::|||:::|||:::|||   :::|||||||||
Db    1602 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1661

Qy     441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
             ||||||||||||||||||:::   |||:::|||:::|||||||||||||||||:::|||
Db    1662 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1721

Qy     461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
             ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1722 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGTAAC 1781

Qy     481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
             ||||||||||||:::||||||||||||||||||||||||||:::||||||::::::
Db    1782 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1841

Qy     501 ValGlnAlaAlaLeuGluLys 507
             ||||||   ||||||:::|||
Db    1842 GTTCAGGTGGCTCTGGATAAG 1862

Search completed: March 31, 2003, 02:24:15
Job time : 7000.72 secs
```

```
OM protein - nucleic search, using frame_plus_p2n model

Run on:          March 30, 2003, 01:20:01 ; Search time 502.349 Seconds
                                            (without alignments)
                                            2304.231 Million cell updates/sec

Title:           US-09-873-409-8
Perfect score:   2592
Sequence:        1 MILGILASLVNGACLPLMPL..........SESKSAVQAALEKDTPRYSF 514

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters:      4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115416_19216/app_query.fasta_1.7544
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_4247_@runat_27032003_115416_19216 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       N_Geneseq_101002:*
                 1:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
                 2:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
                 3:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
                 4:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
                 5:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
                 6:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
                 7:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
                 8:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
                 9:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
                 10:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
                 11:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
                 12:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
                 13:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
                 14:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
                 15:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
                 16:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
                 17:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
                 18:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
                 19:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
                 20:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
                 21:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
                 22:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
                 23:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
                 24:  /SIDS1/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2460.5	94.9	3699	24	ABK83223	Human transporter
2	1590	61.3	4189	21	AAZ49334	Murine multidrug r
3	1590	61.3	4189	24	ABA94367	Mouse BCRP DNA rel
4	1590	61.3	4313	14	AAQ38950	Mouse multidrug re
5	1585	61.1	4279	22	AAD03506	Dog P-glycoprotein
6	1581	61.0	4279	22	AAD03504	Dog P-glycoprotein
7	1581	61.0	4279	22	AAD03505	Dog P-glycoprotein
8	1580.5	61.0	4195	22	AAF86128	Cynomologous monke
9	1579	60.9	3860	21	AAZ49332	Human wild-type mu
10	1579	60.9	3860	24	ABA94365	Human BCRP DNA rel
11	1576	60.8	4233	21	AAZ90198	Rat mdr1b2 (multis
12	1576	60.8	4233	22	AAF27498	Rat mdr1b2 multidr
13	1576	60.8	4279	22	AAD03488	Dog P-glycoprotein
14	1576	60.8	4317	22	AAD03489	Dog P-glycoprotein
15	1576	60.8	4669	19	AAV32645	Human P glycoprote
16	1576	60.8	4669	24	ABK52041	cDNA encoding huma
17	1572	60.6	4186	22	AAF86127	Cynomologous monke
18	1571.5	60.6	4254	24	ABK63517	Rat sequence diffe
19	1570	60.6	3860	21	AAZ49333	Human G185V mutant
20	1570	60.6	3860	24	ABA94366	Human BCRP DNA rel
21	1570	60.6	4349	22	AAH57442	Human intestine ce
22	1570	60.6	4646	15	AAQ72872	Human multidrug re
23	1570	60.6	4646	21	AAZ94738	Human ATP binding
24	1570	60.6	4646	24	AAD38994	Human mdr1 gene.
25	1570	60.6	4646	24	ABL68592	Kidney cancer rela
26	1570	60.6	4646	24	ABL68880	Kidney cancer rela
27	1570	60.6	6505	17	AAT13394	Hybrid vector pSF-
28	1570	60.6	8630	21	AAZ24041	Retroviral M4 mdr-
29	1570	60.6	8630	21	AAZ24042	Retroviral vector
30	1568.5	60.5	4264	19	AAV66533	Mutated human P-gl
31	1568.5	60.5	4264	19	AAV66534	Mutated human P-gl
32	1568	60.5	4669	8	AAN70752	Sequence of human
33	1567	60.5	4378	11	AAQ04522	Multidrug Resistan
34	1566.5	60.4	4425	21	AAZ52048	Rat multidrug resi
35	1564.5	60.4	4369	21	AAZ52047	Rat multidrug resi
36	1564	60.3	4669	14	AAQ52726	Sequence of human
37	1562.5	60.3	3840	24	ABL91687	Human polynucleoti
38	1562.5	60.3	3988	21	AAZ88973	Human MDR-1 DNA.
39	1559.5	60.2	4788	21	AAZ49335	Murine multidrug r
40	1559.5	60.2	4788	24	ABA94368	Mouse BCRP DNA rel
41	1479	57.1	3924	21	AAZ94742	Human ATP binding
42	1479	57.1	3924	21	AAZ88974	Human MDR-3 DNA.
43	1479	57.1	3924	24	ABN95801	Gene #2299 used to
44	1445	55.7	3912	24	ABK63653	Rat sequence diffe
45	1273.5	49.1	4776	21	AAZ94744	Human ATP binding

ALIGNMENTS

```
RESULT 1
ABK83223
ID   ABK83223 standard; cDNA; 3699 BP.
XX
AC   ABK83223;
XX
DT   27-AUG-2002  (first entry)
XX
DE   Human transporter and ion channel, TRICH14, Incyte ID 7472030CB1, cDNA.
XX
KW   Human; ss; gene; transporter and ion channel; TRICH; transport disorder;
KW   neurological disorder; muscle disorder; immunological disorder; cancer;
KW   scleroderma; systemic lupus erythematosus; allergy; leukaemia;
KW   cell proliferative disorder; cervical cancer; breast cancer;
KW   neurodegenerative disorder; Parkinson's disease; Alzheimer's disease;
KW   myotonic dystrophy; catatonia; endocrine disorder; diabetes;
KW   Grave's disease; gastrointestinal disorder; Crohn's disease;
KW   renal disorder; Good pasture's syndrome; viral infection; cirrhosis;
```

```
KW   bacterial infection; fungal infection; parasitic infection;
KW   protozoal infection; helminthic infection; cardiovascular disorder;
KW   atherosclerosis; hepatic disease.
XX
OS   Homo sapiens.
XX
PN   WO200240541-A2.
XX
PD   23-MAY-2002.
XX
PF   25-OCT-2001; 2001WO-US46055.
XX
PR   27-OCT-2000; 2000US-243989P.
PR   03-NOV-2000; 2000US-245904P.
PR   09-NOV-2000; 2000US-247673P.
PR   17-NOV-2000; 2000US-249661P.
PR   20-NOV-2000; 2000US-252232P.
PR   01-DEC-2000; 2000US-250790P.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Tang YT,  Yue H,  Nguyen DB,  Hafalia AJA,  Elliott VS,  Lu Y;
PI   Walia NK,  Yao MG,  Baughn MR,  Gandhi AR,  Ding L,  Sanjanwala M;
PI   Ramkumar J,  Arvizu C,  Gietzen KJ,  Lal PG,  Azimzai Y,  Khan FA;
PI   Thangavelu K,  Thornton M,  Lu DAM,  Tribouley CM,  Warren BA;
PI   Ison CH,  Das D,  Raumann BE,  Policky JL,  Kearney L;
XX
DR   WPI; 2002-463570/49.
DR   P-PSDB; ABG61544.
XX
PT   New transporters and ion channels (TRICH) polypeptides, useful for
PT   diagnosing, preventing, and treating disorders associated with an
PT   abnormal expression or activity of TRICH, e.g. immunological, muscular
PT   or renal disorders  -
XX
PS   Claim 5; Page 172; 178pp; English.
XX
CC   The invention relates to human transporters and ion channels (TRICH)
CC   polypeptides, a naturally occurring amino acid sequence 90 % identical to
CC   TRICH, a biologically active fragment of TRICH or an immunogenic fragment
CC   of TRICH. Also included are an isolated polynucleotide encoding TRICH,
CC   a recombinant polynucleotide comprising a promoter sequence operably
CC   linked to the TRICH polynucleotide, a cell transformed with the
CC   recombinant polynucleotide, a transgenic organism comprising the
CC   recombinant polynucleotide, an isolated antibody that binds specifically
CC   to TRICH, and screening for compounds which bind to TRICH, modulate
CC   TRICH, modulate TRICH expression or are ant/agonists of TRICH.
CC   The polypeptides are useful for diagnosing, treating, and
CC   preventing transport, neurological, muscle, immunological disorders
CC   (e.g. scleroderma, systemic lupus erythematosus, allergies), cell
CC   proliferative disorders such as cancers (e.g. leukaemia, cervical or
CC   breast cancers), neurodegenerative disorders (e.g. Parkinson's disease,
CC   Alzheimer's disease), muscular disorders (e.g. myotonic dystrophy,
CC   catatonia), endocrine disorders (e.g. diabetes, Grave's disease),
CC   gastrointestinal disorders (e.g. Crohn's disease), renal disorders
CC   (e.g. Good pasture's syndrome), viral, bacterial, fungal, parasitic,
CC   protozoal and helminthic infections, cardiovascular disorders (e.g.
CC   atherosclerosis), or hepatic diseases (e.g. cirrhosis) and many
CC   other diseases and disorders detailed in the specification. They can also
CC   be used in assessing the effects of exogenous compounds on the
CC   expression of nucleic acid and amino acid sequences of transporters and
CC   ion channels. TRICH or its fragments may also be used in screening for
CC   compounds that specifically bind to and modulate the activity of TRICH.
CC   The polynucleotides can be used to create knock-in humanised animals or
CC   transgenic animals to model human disease. The present sequence
CC   encodes a TRICH protein.
XX
SQ   Sequence 3699 BP; 1116 A; 707 C; 860 G; 1016 T; 0 other;

Alignment Scores:
Pred. No.:             1.53e-246     Length:         3699
Score:                 2460.50       Matches:        495
Percent Similarity:    95.40%        Conservative:   3

Best Local Similarity:   94.83%        Mismatches:    9
Query Match:             94.93%        Indels:        15
DB:                      24            Gaps:          2

US-09-873-409-8 (1-514) x ABK83223 (1-3699)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          |||||||||||||||||    ||||||    ||||||||||||||||||||||||||||
Db    145 ATGATCCTGGGTATACTGACATCACTGTTCAATGGAGCCTGCCTTCCTTTAATGCCACTG 204

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
             :::||||||||||||||||||||||||||||||||||||||    ||||||||
Db    205 TGTATAGGAGAAATGAGTGATAACCTTATTAGTGGATGTCTAGTCCACACTAACACAACA 264

Qy     41 SerPhePheArg---------------------------------------LeuThr 46
          :::::                                              |||||
Db    265 AATTATCAGAACTGTACTCAGTCTCAAGAGAAGCTGAATGAAGATATGACTCTGTTGACC 324

Qy     47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    325 CTGTATTATGTTGGAATAGGTGTTGCTGCCTTGATTTTTGGTTACATACAGATTTCCTTG 384

Qy     67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    385 TGGATTATAACTGCAGCACGACAGACCAAGAGGATTCGAAAACAGTTTTTTCATTCAGTT 444

Qy     87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    445 TTGGCACAGGACATCGGCTGGTTTGATAGCTGTGACATCGGTGAACTTAACACTCGCATG 504

Qy    107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
          |||   ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    505 ACAGATGACATTGACAAAATCAGTGATGGTATTGGAGATAAGATTGCTCTGTTGTTTCAA 564

Qy    126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    565 AACATGTCTACTTTTTCGATTGGCCTGGCAGTTGGTTTGGTGAAGGGCTGGAAACTCACC 624

Qy    146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    625 CTAGTGACTCTATCCACGTCTCCTCTTATAATGGCTTCAGCGGCAGCATGTTCTAGGATG 684

Qy    166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    685 GTCATCTCATTGACCAGTAAGGAATTAAGTGCCTATTCCAAAGCTGGGGCTGTGGCAGAA 744

Qy    186 GluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGln 205
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    745 GAAGTCTTGTCATCAATCCGAACAGTCATAGCCTTTAGGGCCCAGGAGAAAGAACTTCAA 804

Qy    206 ArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSer 225
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    805 AGGTATACACAGAATCTCAAAGATGCAAAGGATTTTGGCATAAAAAGGACTATAGCTTCA 864

Qy    226 LysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrp 245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    865 AAAGTGTCTCTTGGTGCTGTGTACTTCTTTATGAATGGAACCTATGGACTTGCTTTTTGG 924

Qy    246 TyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAla 265
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    925 TATGGAACCTCCTTGATTCTTAATGGAGAACCTGGATATACCATCGGGACTGTTCTTGCT 984

Qy    266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    985 GTTTTCTTTAGTGTAATCCATAGTAGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAA 1044

Qy    286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1045 ACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTTATTGATAAGAAACCC 1104

Qy    306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
          |||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1105 AGTATAGGTAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGAA 1164
```

```
Qy    326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1165 TTTAAAAATGTTTCTTTCAATTATCCATCAAGGCCATCTATCAAGATTCTGAAAGGTCTG 1224

Qy    346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
          ||||||   |||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1225 AATCTCGGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAG 1284

Qy    366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1285 AGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGTG 1344

Qy    386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1345 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTGGTT 1404

Qy    406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1405 AGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGAGAT 1464

Qy    426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1465 GATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATC 1524

Qy    446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1525 ATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGA 1584

Qy    466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1585 GGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATT 1644

Qy    486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeu 505
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1645 TTAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTCAAGCTGCACTG 1704

Qy    506 GluLys 507
          ||||||
Db   1705 GAGAAG 1710
```

```
RESULT 2
AAZ49334
ID   AAZ49334 standard; cDNA; 4189 BP.
XX
AC   AAZ49334;
XX
DT   14-MAR-2000  (first entry)
XX
DE   Murine multidrug resistance-1 (MDR-1) cDNA.
XX
KW   Multidrug resistance; MDR-1; P-glycoprotein;
KW   transmembrane efflux pump; haematopoietic stem cell; transduction;
KW   bone marrow transplantation; chemotherapy; radiation therapy; cancer;
KW   gene therapy; gene replacement; genetic defect; thalassaemia;
KW   Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
KW   cytokine; ds.
XX
OS   Mus sp.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3831
FT                   /*tag=  a
FT                   /product= "Murine MDR-1 protein"
XX
PN   WO9961589-A2.
XX
PD   02-DEC-1999.
XX
PF   27-MAY-1999;   99WO-US11825.
XX
PR   28-MAY-1998;   98US-0086988.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI   Sorrentino B,  Bunting K;
XX
DR   WPI; 2000-072615/06.
DR   P-PSDB; AAY58188.
XX
PT   Ex vivo expansion of hematopoietic stem cells transduced with a
PT   sequence encoding human multidrug resistance-1, used for bone marrow
PT   transplantation -
XX
PS   Disclosure; Page 90-93; 113pp; English.
XX
CC   This sequence represents cDNA encoding murine multidrug
CC   resistance protein MDR-1. MDR-1 is a transmembrane
CC   efflux pump, responsible for the export of drugs from cells,
CC   particularly cancer cells. The invention relates to transducing
CC   haematopoietic stem cells with nucleic acid encoding an MDR protein
CC   and culturing the modified cells. The modified haematopoietic stem
CC   cells are useful in bone marrow transplantation (to reconstitute
CC   haematopoietic systems in patients who have undergone chemotherapy or
CC   radiation therapy) and in ex vivo gene therapy of genetic defects in
CC   cells derived from haematopoietic stem cells, e.g., thalassaemia,
CC   Gauchier's disease, sickle cell anaemia or leukaemia. The modified
CC   cells can also be used to identify factors involved in regulating
CC   proliferation and differentiation in haematopoietic stem cells.
CC   Haematopoietic stem cells that express MDR-1 will be protected against
CC   chemotherapeutic agents, so can be engrafted while the patient is
CC   undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells
CC   provides sufficient cells to permit standard biochemical analysis.
CC   Overexpression of MDR-1 allows cytokine-driven expansion of
CC   haematopoietic stem cells by at least 10-fold compared with a maximum
CC   of 4-fold in known procedures.
XX
SQ   Sequence 4189 BP; 1204 A; 875 C; 1028 G; 1082 T; 0 other;

Alignment Scores:
Pred. No.:              1.74e-155     Length:         4189
Score:                  1590.00       Matches:        307
Percent Similarity:     76.32%        Conservative:   99
Best Local Similarity:  57.71%        Mismatches:     94
Query Match:            61.34%        Indels:         32
DB:                     21            Gaps:           5

US-09-873-409-8 (1-514) x AAZ49334 (1-4189)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          ||||||||||||   ||||||:::::::::::|||         |||||||||:::   |||
Db    148 ATGATTCTGGGAACTCTCGCTGCTATTATCCATGGAACATTACTTCCCCTCTTGATGCTG 207

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
          |||   |||   |||:::|||:::
Db    208 GTGTTTGGAAACATGACAGATAGTTTTACAAAAGCAGAAGCCAGTATTCTGCCAAGCATT 267

Qy     31 ------------------SerGlyCysLeuValGlnThrAsnThr--------------- 39
                            :::      |||:::   :::|||:::
Db    268 ACTAATCAAAGTGGACCCAACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAAGAG 327

Qy     40 ---------TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56
                   |||::::::                     ||||||   |||||||||
Db    328 ATGGCCATATACGCCTAC---------------TATTACACCGGGATTGGTGCTGGTGTG 372

Qy     57 LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys 76
          ||||||      |||||||||:::|||||||||   :::   |||   ||||||
Db    373 CTCATAGTTGCCTACATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGAAGACAGATACAC 432

Qy     77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96
          :::||||||::::::||||||||:::::::::   |||:::||||||||||||||||
Db    433 AAGATTAGGCAGAAGTTTTTCCATGCTATAATGAATCAGGAGATAGGCTGGTTTGATGTG 492

Qy     97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
             |||:::||||||||||||||||||:::|||   |||:::   ||||||:::||||||
Db    493 CATGATGTTGGGGAGCTCAACACCCGGCTCACAGATGATGTCTCCAAAATTAATGACGGA 552
```

```
Qy     116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
Db     553 ATTGGTGACAAAATTGGGATGTTTTTCAGTCCATAACCACATTTTAGCCGGTTTTATC 612

Qy     136 ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
Db     613 ATAGGATTTATAAGTGGTTGGAAGCTAACCCTTGTCATTTTGGCTGTCAGCCCTCTTATT 672

Qy     156 MetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
Db     673 GGATTGTCATCTGCTTTGTGGGCAAAGGTATTGACTTCATTTACTAATAAGGAACTCCAG 732

Qy     176 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIle 195
Db     733 GCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT 792

Qy     196 AlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLys 215
Db     793 GCCTTTGGAGGACAACAGAAGGAACTTGAAAGGTACAATAAAAATTTAGAAGAAGCTAAA 852

Qy     216 AspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePhe 235
Db     853 AATGTTGGCATAAAGAAAGCTATCACAGCCAGCATTTCGATAGGCATTGCCTACCTGTTG 912

Qy     236 MetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGlu 255
Db     913 GTCTATGCATCATATGCACTGGCATTCTGGTATGGGACATCCTTGGTCCTCTCAAATGAA 972

Qy     256 ProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyr 275
Db     973 ------TATTCTATTGGAGAAGTGCTTACTGTCTTCTTCTCTATTTTGTTGGGGACTTTT 1026

Qy     276 CysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPhe 295
Db    1027 AGTATTGGACACTTGGCCCCAAACATAGAAGCCTTTGCAAACGCACGAGGGGCAGCCTTT 1086

Qy     296 HisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyr 315
Db    1087 GAAATCTTCAAGATAATTGATAACGAGCCAAGCATTGACAGCTTCTCAACAAAGGGCTAC 1146

Qy     316 LysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSer 335
Db    1147 AAACCAGACAGTATAATGGGAAACTTAGAGTTTAAAAATGTTCACTTCAACTACCCATCG 1206

Qy     336 ArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrVal 355
Db    1207 AGAAGCGAAGTTCAGATCTTGAAGGGCCTCAATCTGAAGGTGAAGAGCGGACAGACGGTG 1266

Qy     356 AlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeu 375
Db    1267 GCCTTGGTTGGCAACAGTGGCTGTGGAAAAAGCACAACTGTCCAGCTGATGCAGAGGCTC 1326

Qy     376 TyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVal 395
Db    1327 TACGACCCCCTGGAGGGCGTGGTCAGTATCGACGGACAAGACATCAGAACCATCAATGTG 1386

Qy     396 ArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThr 415
Db    1387 AGGTATCTGAGGGAGATCATTGGTGTGGTGAGTCAGGAACCTGTGCTGTTTGCCACCACG 1446

Qy     416 IleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAla 435
Db    1447 ATCGCCGAGAACATTCGCTATGGCCGAGAAGATGTCACCATGGATGAGATTGAGAAAGCT 1506

Qy     436 AlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeu 455
Db    1507 GTCAAGGAAGCCAATGCCTATGACTTCATCATGAAACTGCCCCACCAATTTGACACCCTG 1566

Qy     456 ValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArg 475
Db    1567 GTTGGTGAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGAATCGCCATTGCCCGG 1626

Qy     476 AlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSer 495
Db    1627 GCCCTGGTCCGCAATCCCAAGATCCTTTTGTTGGACGAGGCCACCTCAGCCCTGGATACA 1686

Qy     496 GluSerLysSerAlaValGlnAlaAlaLeuGluLys 507
Db    1687 GAAAGTGAAGCTGTGGTGCAGGCCGCACTGGATAAG 1722
```

```
RESULT 3
ABA94367
ID   ABA94367 standard; DNA; 4189 BP.
XX
AC   ABA94367;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Mouse BCRP DNA related seq Id No. 5.
XX
KW   Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;
KW   hepatotropic; neurodegenerative; cytostatic; antianemic; muscular; BCRP;
KW   cardiant; gene therapy; ds.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3831
FT                   /*tag=  a
XX
PN   WO200192877-A2.
XX
PD   06-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US17459.
XX
PR   31-MAY-2000; 2000US-0584586.
PR   29-MAY-2001; 2001US-0866866.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI   Sorrentino B,  Schuetz J;
XX
DR   WPI; 2002-114368/15.
DR   P-PSDB; ABB07268.
XX
PT   Identifying a stem cell, for treating e.g., muscular dystrophy,
PT   myocardial infarction, Parkinson's disease, or neurodegenerative
PT   disorders, comprises detecting the expression of an ATP transport
PT   protein (BCRP) by a cell  -
XX
PS   Disclosure; Page 64-66; 87pp; English.
XX
CC   The invention provides a method of identifying and/or isolating a stem
CC   cell that involves detecting the expression of an ATP transport protein
CC   containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
CC   comprising stem cells. The isolated stem cells may be used in the
CC   treatment of diseases such as muscular dystrophy, degenerative liver
CC   disorder, myocardial infarction, Parkinson's disease, degenerative
CC   disorders of the brain, and for tissue regeneration or replacement.
CC   Haematopoietic cells can be used in bone marrow transplants (e.g., for
CC   treatment of leukemia) and for ex vivo gene therapy for treating blood
CC   diseases such as sickle cell anemia and thalassemia. The stem cells can
CC   also be used as cell targets in gene therapy protocols. The present
CC   sequence represents a sequence related to the BCRP for which no relevant
CC   information has been provided in the specification.
XX
SQ   Sequence 4189 BP; 1204 A; 875 C; 1028 G; 1082 T; 0 other;

Alignment Scores:
Pred. No.:              1.74e-155     Length:         4189
Score:                  1590.00       Matches:        307
Percent Similarity:     76.32%        Conservative:   99
Best Local Similarity:  57.71%        Mismatches:     94
Query Match:            61.34%        Indels:         32
```

```
DB:                      24             Gaps:        5

US-09-873-409-8 (1-514) x ABA94367 (1-4189)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            |||||||||||    |||||::::::::::::|||        ||||||||:::    |||
Db      148 ATGATTCTGGGAACTCTCGCTGCTATTATCCATGGAACATTACTTCCCCTCTTGATGCTG 207

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
            |||   |||   |||:::|||:::
Db      208 GTGTTTGGAAACATGACAGATAGTTTTACAAAAGCAGAAGCCAGTATTCTGCCAAGCATT 267

Qy       31 ------------------SerGlyCysLeuValGlnThrAsnThr--------------- 39
                              :::      |||:::   :::|||:::
Db      268 ACTAATCAAAGTGGACCCAACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAAGAG 327

Qy       40 ---------TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56
                     |||::::::               ||||||  ||||||||
Db      328 ATGGCCATATACGCCTAC---------------TATTACACCGGGATTGGTGCTGGTGTG 372

Qy       57 LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys 76
            ||||||      |||||||||:::||||||||    :::   |||   ||||||
Db      373 CTCATAGTTGCCTACATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGAAGACAGATACAC 432

Qy       77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96
            :::||||||::::::|||||||||:::::::::   |||:::|||||||||||||||
Db      433 AAGATTAGGCAGAAGTTTTTCCATGCTATAATGAATCAGGAGATAGGCTGGTTTGATGTG 492

Qy       97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
               |||:::||||||||||||||||||:::|||   |||:::   ||||||:::||||||
Db      493 CATGATGTTGGGGAGCTCAACACCCGGCTCACAGATGATGTCTCCAAAATTAATGACGGA 552

Qy      116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
            ||||||||||||||    :::   ||||||:::::::::||||||      |||
Db      553 ATTGGTGACAAAATTGGGATGTTTTTTCAGTCCATAACCACATTTTTAGCCGGTTTTATC 612

Qy      136 ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
            :::|||   :::   ||||||||||||||||||||||||:::   ||||||||||||
Db      613 ATAGGATTTATAAGTGGTTGGAAGCTAACCCTTGTCATTTTGGCTGTCAGCCCTCTTATT 672

Qy      156 MetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
                     |||:::|||    ::::::::::::   |||   |||:::|||||||||
Db      673 GGATTGTCATCTGCTTTGTGGGCAAAGGTATTGACTTCATTTACTAATAAGGAACTCCAG 732

Qy      176 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIle 195
            ||||||:::|||||||||||||||||||||||||||||||::::::|||||||||||||||
Db      733 GCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT 792

Qy      196 AlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLys 215
            ||||||      |||:::|||||||||:::||||||   :::||||||::::::||||||
Db      793 GCCTTTGGAGGACAACAGAAGGAACTTGAAAGGTACAATAAAAATTTAGAAGAAGCTAAA 852

Qy      216 AspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePhe 235
            :::   |||||||||:::   |||   :::   :::|||:::|||      |||
Db      853 AATGTTGGCATAAAGAAAGCTATCACAGCCAGCATTTCGATAGGCATTGCCTACCTGTTG 912

Qy      236 MetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGlu 255
            :::      :::|||   |||||||||||||||||||||||||||:::|||:::   |||
Db      913 GTCTATGCATCATATGCACTGGCATTCTGGTATGGGACATCCTTGGTCCTCTCAAATGAA 972

Qy      256 ProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyr 275
                     |||:::||||||   ||||||   ||||||||||||::::::       ::::::
Db      973 ------TATTCTATTGGAGAAGTGCTTACTGTCTTCTTCTCTATTTTGTTGGGGACTTTT 1026

Qy      276 CysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPhe 295
               ||||||            |||:::   |||   ||||||   ||||||||||||||||||
Db     1027 AGTATTGGACACTTGGCCCCAAACATAGAAGCCTTTGCAAACGCACGAGGGGCAGCCTTT 1086

Qy      296 HisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyr 315
               ||||||:::::|||||      :::||||||||||||:::||||||||||      ||||||
Db     1087 GAAATCTTCAAGATAATTGATAACGAGCCAAGCATTGACAGCTTCTCAACAAAGGGCTAC 1146

Qy      316 LysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSer 335
```

```
            ||||||:::||||||   |||   :::||||||||||||||    ||||||||||||||
Db     1147 AAACCAGACAGTATAATGGGAAACTTAGAGTTTAAAAATGTTCACTTCAACTACCCATCG 1206

Qy      336 ArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrVal 355
            |||      ::::::||||||||||||||||||||||||::::::|||||||||:::||||||
Db     1207 AGAAGCGAAGTTCAGATCTTGAAGGGCCTCAATCTGAAGGTGAAGAGCGGACAGACGGTG 1266

Qy      356 AlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeu 375
            ||||||||||||    :::|||   ||||||||||||   |||||||||:::||||||||
Db     1267 GCCTTGGTTGGCAACAGTGGCTGTGGAAAAAGCACAACTGTCCAGCTGATGCAGAGGCTC 1326

Qy      376 TyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVal 395
            |||||||||   :::|||   :::   :::|||      |||||||||   :::||||||
Db     1327 TACGACCCCCTGGAGGGCGTGGTCAGTATCGACGGACAAGACATCAGAACCATCAATGTG 1386

Qy      396 ArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThr 415
            |||:::   |||:::   |||||||||||||||||||||||||||||||||   ||||||
Db     1387 AGGTATCTGAGGGAGATCATTGGTGTGGTGAGTCAGGAACCTGTGCTGTTTGCCACCACG 1446

Qy      416 IleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAla 435
            |||:::   ||||||:::||||||||:::|||||||||   :::|||:::|||:::|||
Db     1447 ATCGCCGAGAACATTCGCTATGGCCGAGAAGATGTCACCATGGATGAGATTGAGAAAGCT 1506

Qy      436 AlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeu 455
               :::||||||||||||||||||||||||:::      |||::::::|||:::||||||
Db     1507 GTCAAGGAAGCCAATGCCTATGACTTCATCATGAAACTGCCCCACCAATTTGACACCCTG 1566

Qy      456 ValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArg 475
            ||||||||:::|||||||||:::|||||||||||||||||||||||||||||||||||
Db     1567 GTTGGTGAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGAATCGCCATTGCCCGG 1626

Qy      476 AlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSer 495
            |||||||||||||||||||||||||:::||||||||||||||||||||||||:::
Db     1627 GCCCTGGTCCGCAATCCCAAGATCCTTTTGTTGGACGAGGCCACCTCAGCCCTGGATACA 1686

Qy      496 GluSerLysSerAlaValGlnAlaAlaLeuGluLys 507
            ||||||::::::   ||||||||||||:::|||
Db     1687 GAAAGTGAAGCTGTGGTGCAGGCCGCACTGGATAAG 1722
```

```
RESULT 4
AAQ38950
ID   AAQ38950 standard; DNA; 4313 BP.
XX
AC   AAQ38950;
XX
DT   28-JUL-1993  (first entry)
XX
DE   Mouse multidrug resistance sequence.
XX
KW   mdr gene; Lambda DR11 clone; ss.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   CDS             110..3940
FT                   /*tag= a
FT                   /phenotype= multidrug_resistance
XX
PN   US5198344-A.
XX
PD   30-MAR-1993.
XX
PF   15-JUL-1986;   86US-0885951.
XX
PR   15-JUL-1986;   86US-0885951.
PR   06-FEB-1991;   91US-0652311.
XX
PA   (MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
PI   Croop JM,  Gros P,  Housman DE;
XX
DR   WPI; 1993-126077/15.
DR   P-PSDB; AAR35199.
```

```
XX
PT   DNA sequence which confers multi-drug resistance on sensitive
PT   mammalian cells - used to preserve bone marrow cells during
PT   chemotherapy to prevent infection
XX
PS   Claim 1; Fig 8; 22pp; English.
XX
CC   A cDNA library was constructed from mRNA which had been isolated
CC   from a drug-sensitive mouse cell-line. Two mouse cDNA molecules
CC   complementary to the mRNA species encoded by 2 related but distinct
CC   mdr genes were isolated and cloned. One of the cDNA clones (lambda
CC   DR11) is a full-length cDNA clone for one member of the mdr gene
CC   family. When incorporated into prokaryotic expression vector pDREX4
CC   (which allows high levels of transcription of the cDNA when
CC   introduced into mammalian cells), the clone was shown to confer the
CC   multidrug resistance phenotype upon transfection into drug-sensitive
CC   mammalian cells.
XX
SQ   Sequence 4313 BP; 1241 A; 904 C; 1058 G; 1110 T; 0 other;

Alignment Scores:
Pred. No.:               1.81e-155      Length:        4313
Score:                   1590.00        Matches:       307
Percent Similarity:      76.32%         Conservative:  99
Best Local Similarity:   57.71%         Mismatches:    94
Query Match:             61.34%         Indels:        32
DB:                      14             Gaps:          5

US-09-873-409-8 (1-514) x AAQ38950 (1-4313)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          |||||||||||    ||||||::::::::::::|||      |||||||||:::    |||
Db    257 ATGATTCTGGGAACTCTCGCTGCTATTATCCATGGAACATTACTTCCCCTCTTGATGCTG 316

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
            |||  |||   |||:::|||:::
Db    317 GTGTTTGGAAACATGACAGATAGTTTACAAAAGCAGAAGCCAGTATTCTGCCAAGCATT 376

Qy     31 ------------------SerGlyCysLeuValGlnThrAsnThr--------------- 39
                            :::      |||:::   :::|||:::
Db    377 ACTAATCAAAGTGGACCCAACAGTACTCTGATCATCAGCAACAGCAGTCTGGAGGAAGAG 436

Qy     40 ---------TyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAla 56
                   |||::::::                |||||   ||||||||
Db    437 ATGGCCATATACGCCTAC---------------TATTACACCGGGATTGGTGCTGGTGTG 481

Qy     57 LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys 76
          ||||||   ||||||||||:::||||||||    :::   |||     ||||||
Db    482 CTCATAGTTGCCTACATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGAAGACAGATACAC 541

Qy     77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96
          :::||||||::::::|||||||||:::::::::   |||:::|||||||||||||||
Db    542 AAGATTAGGCAGAAGTTTTTCCATGCTATAATGAATCAGGAGATAGGCTGGTTTGATGTG 601

Qy     97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
                |||:::|||||||||||||||||:::|||   |||:::   ||||||:::||||||
Db    602 CATGATGTTGGGGAGCTCAACACCCGGCTCACAGATGATGTCTCCAAAATTAATGACGGA 661

Qy    116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
          |||||||||||||||   :::   ||||||:::::::::||||||         |||
Db    662 ATTGGTGACAAAATTGGGATGTTTTTCAGTCCATAACCACATTTTAGCCGGTTTTATC 721

Qy    136 ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
          :::|||   :::   |||||||||||||||||||||   |||:::   ||||||||||||
Db    722 ATAGGATTTATAAGTGGTTGGAAGCTAACCCTTGTCATTTTGGCTGTCAGCCCTCTTATT 781

Qy    156 MetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
                |||:::|||                 |||   |||:::|||||||||
Db    782 GGATTGTCATCTGCTTTGTGGGCAAAGGTATTGACTTCATTTACTAATAAGGAACTCCAG 841

Qy    176 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIle 195
          ||||||:::|||||||||||||||||||||||||||||::::::||||||||||||||||
Db    842 GCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTCTTAGCAGCCATCAGAACTGTGATT 901

Qy    196 AlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLys 215
          ||||||      |||:::|||||||||:::||||||      :::||||||:::::|||||
Db    902 GCCTTTGGAGGACAACAGAAGGAACTTGAAAGGTACAATAAAAATTTAGAAGAAGCTAAA 961

Qy    216 AspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePhe 235
          :::   |||||||||:::   |||   :::   :::|||:::|||      |||
Db    962 AATGTTGGCATAAAGAAAGCTATCACAGCCAGCATTTCGATAGGCATTGCCTACCTGTTG 1021

Qy    236 MetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGlu 255
          :::      :::|||   |||||||||||||||||||||||||:::|||:::   |||
Db   1022 GTCTATGCATCATATGCACTGGCATTCTGGTATGGGACATCCTTGGTCCTCTCAAATGAA 1081

Qy    256 ProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyr 275
                |||:::||||||   ||||||   |||||||||||||::::::      ::::::
Db   1082 ------TATTCTATTGGAGAAGTGCTTACTGTCTTCTTCTCTATTTTGTTGGGGACTTTT 1135

Qy    276 CysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPhe 295
                ||||||         |||:::   |||   ||||||   ||||||||||||||||||
Db   1136 AGTATTGGACACTTGGCCCCAAACATAGAAGCCTTTGCAAACGCACGAGGGGCAGCCTTT 1195

Qy    296 HisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyr 315
                ||||||::::::||||||   :::||||||||||||:::|||||||||   ||||||
Db   1196 GAAATCTTCAAGATAATTGATAACGAGCCAAGCATTGACAGCTTCTCAACAAAGGGCTAC 1255

Qy    316 LysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSer 335
          ||||||:::||||||   |||   :::||||||||||||||   |||||||||||||||
Db   1256 AAACCAGACAGTATAATGGGAAACTTAGAGTTTAAAAATGTTCACTTCAACTACCCATCG 1315

Qy    336 ArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrVal 355
          |||      ::::::||||||||||||||||||||||||::::::|||||||||:::||||||
Db   1316 AGAAGCGAAGTTCAGATCTTGAAGGGCCTCAATCTGAAGGTGAAGAGCGGACAGACGGTG 1375

Qy    356 AlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeu 375
          ||||||||||||      :::|||   ||||||||||||   ||||||||::::||||||||
Db   1376 GCCTTGGTTGGCAACAGTGGCTGTGGAAAAAGCACAACTGTCCAGCTGATGCAGAGGCTC 1435

Qy    376 TyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVal 395
          ||||||||   :::|||   :::   :::|||      ||||||||||   :::||||||
Db   1436 TACGACCCCCTGGAGGGCGTGGTCAGTATCGACGGACAAGACATCAGAACCATCAATGTG 1495

Qy    396 ArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThr 415
          |||:::   |||:::   |||||||||||||||||||||||||||||||   ||||||
Db   1496 AGGTATCTGAGGGAGATCATTGGTGTGGTGAGTCAGGAACCTGTGCTGTTTGCCACCACG 1555

Qy    416 IleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAla 435
          |||:::   ||||||:::||||||||||:::|||||||||   :::|||:::|||:::|||
Db   1556 ATCGCCGAGAACATTCGCTATGGCCGAGAAGATGTCACCATGGATGAGATTGAGAAAGCT 1615

Qy    436 AlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeu 455
             :::||||||||||||||||||||||||||||:::   |||::::::|||:::||||||
Db   1616 GTCAAGGAAGCCAATGCCTATGACTTCATCATGAAACTGCCCCACCAATTTGACACCCTG 1675

Qy    456 ValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArg 475
          |||||||||:::||||||||:::||||||||||||||||||||||||||||||||||||
Db   1676 GTTGGTGAGAGAGGGGCGCAGCTGAGTGGGGGACAGAAACAGAGAATCGCCATTGCCCGG 1735

Qy    476 AlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSer 495
          ||||||||||||||||||||||||:::||||||||||||||||||||||||||||:::
Db   1736 GCCCTGGTCCGCAATCCCAAGATCCTTTTGTTGGACGAGGCCACCTCAGCCCTGGATACA 1795

Qy    496 GluSerLysSerAlaValGlnAlaAlaLeuGluLys 507
          ||||||:::::::   ||||||||||||:::|||
Db   1796 GAAAGTGAAGCTGTGGTGCAGGCCGCACTGGATAAG 1831

RESULT 5
AAD03506
ID   AAD03506 standard; cDNA; 4279 BP.
XX
AC   AAD03506;
XX
DT   13-JUN-2001  (first entry)
XX
```

```
DE   Dog P-glycoprotein (PGP) allelic variant (Genotype D) cDNA.
XX
KW   Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KW   MDR1; drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag= a
FT                   /product= "Dog P-glycoprotein (PGP) allelic variant
FT                   (Genotype D) protein"
FT   allele          replace (91, T)
FT                   /*tag= b
FT   allele          replace (607, C)
FT                   /*tag= c
FT   allele          replace (1001, T)
FT                   /*tag= c
FT   allele          replace (3458, A)
FT                   /*tag= c
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00310.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 9; Page 102-107; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) allelic variant
CC   (Genotype D) cDNA. The PGP enzyme functions as an efflux pump
CC   exporting small molecules across the cell membrane. This enzyme
CC   is a member of the ABC transporter family.
XX
SQ   Sequence 4279 BP; 1296 A; 833 C; 1009 G; 1141 T; 0 other;

Alignment Scores:
Pred. No.:              5.96e-155      Length:        4279
Score:                  1585.00        Matches:       304
Percent Similarity:     75.80%         Conservative:  97
Best Local Similarity:  57.47%         Mismatches:    104
Query Match:            61.15%         Indels:        24
DB:                     22             Gaps:          4

US-09-873-409-8 (1-514) x AAD03506 (1-4279)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
Db    227 GTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy     35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
Db    287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy     45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
Db    347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy     60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
Db    407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy     80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
Db    467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy    100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
Db    527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy    119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
Db    587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy    139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
Db    647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy    159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
Db    707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy    179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
Db    767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy    199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
Db    827 GGACAAAAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAGGAATTGGG 886

Qy    219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238
Db    887 ATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCA 946

Qy    239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
Db    947 TCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TAT 1000

Qy    259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
Db   1001 ACTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGA 1060

Qy    279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
Db   1061 CAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTC 1120

Qy    299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
Db   1121 AAGATAATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGAT 1180

Qy    319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
Db   1181 AATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAA 1240

Qy    339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
```

```
Db    1241 GTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTT 1300

Qy     359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
           |||   :::|||   |||||||||||   ||||||||:::|||||||||||||||
Db    1301 GGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCC 1360

Qy     379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
              |||||   :::   :::|||      ||||||||   :::|||||||||||
Db    1361 ACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTT 1420

Qy     399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
           |||:::      ||||||||||||||||||||||||||||   |||||||||:::
Db    1421 CGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAA 1480

Qy     419 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
           ||||||:::|||||||||::::::||||||   :::|||:::|||:::|||   :::|||
Db    1481 AACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAA 1540

Qy     439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
           ||||||||||||||||||||||||:::   ||||||||||:::|||||||||||||||
Db    1541 GCCAATGCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAG 1600

Qy     459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
           :::|||||||||:::||||||||||||||||||||||||||||||||||||||||||
Db    1601 AGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTT 1660

Qy     479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
           |||||||||||||||||:::||||||||||||||||||||||||:::||||||:::
Db    1661 CGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAA 1720

Qy     499 SerAlaValGlnAlaAlaLeuGluLys 507
           :::   ||||||   ||||||:::|||
Db    1721 GCAGTGGTTCAGGTGGCCCTGGATAAG 1747

RESULT 6
AAD03504
ID   AAD03504 standard; cDNA; 4279 BP.
XX
AC   AAD03504;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) allelic variant (Genotype A) cDNA.
XX
KW   Dog; P-glycoprotein allelic variant; multidrug transporter; MDR1;
KW   drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag=  a
FT                   /product= "Dog P-glycoprotein (PGP) allelic variant
FT                   (Genotype A) protein"
FT   allele          replace (607, C)
FT                   /*tag=  b
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00308.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 9; Page 85-90; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) allelic variant
CC   (Genotype A) cDNA. The PGP enzyme functions as an efflux pump
CC   exporting small molecules across the cell membrane. This enzyme
CC   is a member of the ABC transporter family.
XX
SQ   Sequence 4279 BP; 1295 A; 833 C; 1008 G; 1143 T; 0 other;

Alignment Scores:
Pred. No.:               1.56e-154     Length:          4279
Score:                   1581.00       Matches:         303
Percent Similarity:      75.80%        Conservative:    98
Best Local Similarity:   57.28%        Mismatches:      104
Query Match:             61.00%        Indels:          24
DB:                      22            Gaps:            4

US-09-873-409-8 (1-514) x AAD03504 (1-4279)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||::::::|||   :::|||::::::::::::||||||   |||||||||||   |||
Db     167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
           |||   |||   |||:::|||:::         :::      :::
Db     227 GTTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy      35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                                :::   ||||||   |||
Db     287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy      45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                             ||||||   |||||||||         |||:::
Db     347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy      60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
              |||||||||:::|||   |||   :::   |||   ||||||      :::||||||
Db     407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy      80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
           |||||||||||||||:::::::::   |||:::||||||||||||||||   |||:::
Db     467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy     100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
           ||||||||||||||||||:::|||   |||:::   ||||||::::::||||||||||||
Db     527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy     119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
           ||||||   :::   ||||||:::::::::||||||      |||      ||||||
Db     587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy     139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
              :::||||||||||||||||||||   |||:::   ||||||::::::      |||
Db     647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy     159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
           ||||||      ::::::::::::   |||   |||   |||||||||   ||||||:::
```

```
Db     707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy     179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
           ||||||||||||||||||||||||||||||::::::|||||||||||||||||||
Db     767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy     199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
              |||:::||||||||:::||||||   :::||||||::::::||||||      |||
Db     827 GGACAAAAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAGGAATTGGG 886

Qy     219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238
           ||||||:::   |||   :::   :::|||:::||||||   :::      :::
Db     887 ATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCA 946

Qy     239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
           :::|||   ||||||||||||||||||||||||||||:::|||:::   |||      |||
Db     947 TCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TAT 1000

Qy     259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
           :::||||||   ||||||   |||||||||||||||:::      ::::::   ||||||
Db    1001 TCTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGA 1060

Qy     279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
              |||   |||         |||   ||||||   ||||||||||||||||:::   ||||||
Db    1061 CAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTC 1120

Qy     299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
           ::::::||||||   ||||||||||||||||||::::::|||   :::|||:::||||||:::
Db    1121 AAGATAATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGAT 1180

Qy     319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
           :::|||:::|||   :::|||||||||||||||   |||:::||||||||||||
Db    1181 AATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAA 1240

Qy     339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
           :::||||||||||||||||||||||||::::::::|||||||:::|||||||||||||||
Db    1241 GTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTT 1300

Qy     359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
           |||   :::|||   ||||||||||||   |||||||||:::||||||||||||||||||
Db    1301 GGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCC 1360

Qy     379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
              ||||||   :::   :::|||      ||||||||:::||||||||||||
Db    1361 ACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTT 1420

Qy     399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
           |||:::      ||||||||||||||||||||||||||||||   ||||||||||:::
Db    1421 CGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAA 1480

Qy     419 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
           ||||||:::||||||||::::::||||||   :::|||:::|||:::|||   :::|||
Db    1481 AACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAA 1540

Qy     439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
           ||||||||||||||||||||||||:::   ||||||||||||:::||||||||||||||||
Db    1541 GCCAATGCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAG 1600

Qy     459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
           :::|||||||||:::||||||||||||||||||||||||||||||||||||||||||||
Db    1601 AGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTT 1660

Qy     479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
           ||||||||||||:::||||||||||||||||||||||||||||:::||||||:::
Db    1661 CGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAA 1720

Qy     499 SerAlaValGlnAlaAlaLeuGluLys 507
           :::   ||||||   |||||||:::|||
Db    1721 GCAGTGGTTCAGGTGGCCCTGGATAAG 1747

RESULT 7
AAD03505
ID   AAD03505 standard; cDNA; 4279 BP.
XX
AC   AAD03505;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) allelic variant (Genotype B) cDNA.
XX
KW   Dog; P-glycoprotein allelic variant; PGP; multidrug transporter;
KW   MDR1; drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag=  a
FT                   /product= "Dog P-glycoprotein (PGP) allelic variant
FT                   (Genotype B) protein"
FT   allele          replace (91, T)
FT                   /*tag=  b
FT   allele          replace (607, C)
FT                   /*tag=  c
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00309.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 9; Page 93-99; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) allelic variant
CC   (Genotype B) cDNA. The PGP enzyme functions as an efflux pump
CC   exporting small molecules across the cell membrane. This enzyme
CC   is a member of the ABC transporter family.
XX
SQ   Sequence 4279 BP; 1296 A; 833 C; 1008 G; 1142 T; 0 other;

Alignment Scores:
Pred. No.:              1.56e-154      Length:         4279
Score:                  1581.00        Matches:        303
Percent Similarity:     75.80%         Conservative:   98
Best Local Similarity:  57.28%         Mismatches:     104
Query Match:            61.00%         Indels:         24
DB:                     22             Gaps:           4

US-09-873-409-8 (1-514) x AAD03505 (1-4279)
```

```
Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||:::::|||   :::|||::::::::::::|||||  ||||||||||   |||
Db     167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
           |||  |||  |||:::|||:::    :::    :::
Db     227 GTTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy      35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                                :::     |||||     |||
Db     287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy      45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                                |||||  ||||||||        |||:::
Db     347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy      60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
              ||||||||:::|||  |||  :::  |||  |||||     :::||||||
Db     407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy      80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
           ||||||||||||||:::::::::  |||:::||||||||||||||||    |||:::
Db     467 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy     100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
           ||||||||||||||||:::|||    |||:::  ||||||:::::::||||||||||
Db     527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy     119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
           ||||||   :::   ||||||:::::::::||||||   |||     ||||||
Db     587 AAAATTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy     139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
              :::||||||||||||||||||||  |||:::  ||||||::::::       |||
Db     647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy     159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
           ||||||   ::::::::::::  |||  |||   |||||||||    ||||||:::
Db     707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy     179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
           ||||||||||||||||||||||||||||::::::||||||||||||||||||||
Db     767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy     199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
              |||:::||||||||:::||||||  :::||||||::::::||||||     |||
Db     827 GGACAAAAGAAAGAACTTGAAAGGTACAACAAAAAATTAGAAGAAGCTAAAGGAATTGGG 886

Qy     219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238
           ||||||:::  |||  :::  :::|||:::||||||  :::      :::
Db     887 ATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCA 946

Qy     239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
           :::|||  ||||||||||||||||||||||||||||:::|||:::   |||      |||
Db     947 TCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TAT 1000

Qy     259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
           :::||||||  ||||||  ||||||||||||||||:::       ::::::    ||||||
Db    1001 TCTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGA 1060

Qy     279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
              |||  |||     |||  ||||||  ||||||||||||||:::   ||||||
Db    1061 CAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTC 1120

Qy     299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
           ::::::||||||   ||||||||||||||||::::::|||   :::|||:::||||||:::
Db    1121 AAGATAATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGAT 1180

Qy     319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
           :::|||:::|||   :::||||||||||||||   |||:::||||||||||||
Db    1181 AATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAA 1240

Qy     339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
           :::||||||||||||||||||||||||:::::::::||||||:::||||||||||||||
Db    1241 GTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTT 1300

Qy     359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
           |||   :::|||   ||||||||||||   |||||||||:::||||||||||||||||
Db    1301 GGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCC 1360

Qy     379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
                 |||||   :::   :::|||       ||||||||   :::||||||||||||
Db    1361 ACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTT 1420

Qy     399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
           |||:::      ||||||||||||||||||||||||||||   |||||||||:::
Db    1421 CGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAA 1480

Qy     419 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
           ||||||:::||||||||||::::::||||||   :::|||:::|||:::|||   :::|||
Db    1481 AACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAA 1540

Qy     439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
           ||||||||||||||||||||||||:::   ||||||||||||:::||||||||||||||||
Db    1541 GCCAATGCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAG 1600

Qy     459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
           :::|||||||||:::|||||||||||||||||||||||||||||||||||||||||||
Db    1601 AGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTT 1660

Qy     479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
           ||||||||||||||||:::||||||||||||||||||||||||||:::||||||:::
Db    1661 CGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAA 1720

Qy     499 SerAlaValGlnAlaAlaLeuGluLys 507
           :::   ||||||  ||||||:::|||
Db    1721 GCAGTGGTTCAGGTGGCCCTGGATAAG 1747
```

```
RESULT 8
AAF86128
ID   AAF86128 standard; cDNA; 4195 BP.
XX
AC   AAF86128;
XX
DT   25-JUN-2001  (first entry)
XX
DE   Cynomologous monkey P-glycoprotein cDNA variant 1.
XX
KW   Cynomologous monkey; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   efflux pump; ss.
XX
OS   Macaca fascicularis.
XX
FH   Key             Location/Qualifiers
FT   CDS             100..3951
FT                   /*tag=  a
FT                   /product= "PGP"
FT                   /note= "P-glycoprotein"
FT   misc_feature    376..384
FT                   /*tag=  b
FT                   /note= "Insertion of 9 nucleotides relative to PGP
FT                          allelic variant AAF86127"
XX
PN   WO200123565-A1.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26592.
XX
PR   28-SEP-1999;   99US-0156921.
PR   12-OCT-1999;   99US-0158818.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-Crespi DT,  Crespi CL;
XX
```

```
DR   WPI; 2001-316136/33.
DR   P-PSDB; AAB81065.
XX
PT   Novel isolated nucleic acid encoding cynomologous monkey P-glycoprotein
PT   (PGP) and homologous PGP polypeptides are useful for predicting
PT   bioavailability of compound and increasing PGP transporter activity in
PT   cell  -
XX
PS   Example 1; Page 59-65; 84pp; English.
XX
CC   This invention relates to a polynucleotide sequence encoding a
CC   cynomologous monkey P-glycoprotein (PGP), and an allelic variant of the
CC   PGP protein. PGP, also known as multidrug transporter, MDR1 is a member
CC   of the ABC transporter superfamily. The enzyme serves as an efflux pump
CC   exporting small molecules across the cell membrane. The invention
CC   includes a cynomologous monkey (Macaca fascicularis) PGP coding sequence
CC   and protein, and also that of an allelic variant. The PGP polynucleotide
CC   sequence is useful for increasing PGP transporter activity in a cell.
CC   Antisense sequences of the cDNA are useful for inhibiting PGP transport
CC   activity in a mammalian cell. They may also be used for increasing the
CC   bioavailability of a drug. The present sequence represents cDNA encoding
CC   cynomologous monkey P-glycoprotein. This sequence contains a 9 nucleotide
CC   insert compared to the PGP allelic variant given in AAF86127.
XX
SQ   Sequence 4195 BP; 1230 A; 801 C; 1039 G; 1125 T; 0 other;

Alignment Scores:
Pred. No.:              1.71e-154     Length:        4195
Score:                  1580.50       Matches:       306
Percent Similarity:     75.85%        Conservative:  96
Best Local Similarity:  57.74%        Mismatches:    103
Query Match:            60.98%        Indels:        25
DB:                     22            Gaps:          4

US-09-873-409-8 (1-514) x AAF86128 (1-4195)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    250 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGAGCTGGACTTCCTCTCATGATGCTG 309

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSer---------------------GlyCys 33
Db    310 GTGTTTGGAGACATGACGGATACCTTTGCAAATGCAGGAAATTTAGGAGATTTAGGAGCT 369

Qy     34 LeuValGlnThrAsnThrTyrSer------------------------------------ 41
Db    370 CTGTTGTTTAACAACACTAATAGCAGTAATATCACTGATACAGTGCCCGTCATGAATCTG 429

Qy     42 ---------PhePheArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIle 58
Db    430 GAGGAAGATATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTT 489

Qy     59 PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIle 78
Db    490 GCTGCTTACATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATT 549

Qy     79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp 98
Db    550 AGAAAACAGTTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGAT 609

Qy     99 IleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGly 117
Db    610 GTTGGGGAGCTTAACACCCGGCTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGT 669

Qy    118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
Db    670 GACAAAATTGGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGA 729

Qy    138 LeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
Db    730 TTTACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTG 789

Qy    158 SerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr 177
Db    790 TCAGCTGCAGTCTGGGCAAAGATACTGTCTTCATTTACTGATAAAGAACTCTTAGCTTAT 849

Qy    178 SerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197
Db    850 GCAAAAGCTGGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTT 909

Qy    198 ArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPhe 217
Db    910 GGAGGACAAAAGAAAGAACTCGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATT 969

Qy    218 GlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsn 237
Db    970 GGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTTATCTAT 1029

Qy    238 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGly 257
Db   1030 GCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAAAGGAA------ 1083

Qy    258 TyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIle 277
Db   1084 TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTT 1143

Qy    278 GlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIle 297
Db   1144 GGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTTTGAAATC 1203

Qy    298 PheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysPro 317
Db   1204 TTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCA 1263

Qy    318 GluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgPro 337
Db   1264 GATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAA 1323

Qy    338 SerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeu 357
Db   1324 GAAGTTAAGATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTG 1383

Qy    358 ValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAsp 377
Db   1384 GTTGGAAACAGCGGCTGTGGGAAGAGCACAACGGTCCAGCTGATGCAGAGGCTTTATGAC 1443

Qy    378 ProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHis 397
Db   1444 CCCACAGAGGGCATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAACGTAAGGTTT 1503

Qy    398 TyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSer 417
Db   1504 CTACGGGAAATCATCGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCT 1563

Qy    418 AsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArg 437
Db   1564 GAAAACATTCGCTATGGTCGTGAAGATGTCACCATGGATGAGATTGAGAAAGCTGTCAAG 1623

Qy    438 GluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGly 457
Db   1624 GAAGCCAATGCCTATGACTTTATCATGAAACTGCCTCAGAAATTTGACACCCTGGTTGGA 1683

Qy    458 GluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeu 477
Db   1684 GAGAGAGGGGCCCAGCTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTG 1743

Qy    478 ValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSer 497
Db   1744 GTTCGCAACCCCAAGATCCTCCTGCTGGACGAGGCCACGTCAGCCTTGGACACAGAAAGT 1803

Qy    498 LysSerAlaValGlnAlaAlaLeuGluLys 507
Db   1804 GAAGCAGTGGTTCAGGTGGCTCTGGATAAG 1833

RESULT 9
AAZ49332
```

```
ID   AAZ49332 standard; cDNA; 3860 BP.
XX
AC   AAZ49332;
XX
DT   14-MAR-2000  (first entry)
XX
DE   Human wild-type multidrug resistance-1 (MDR-1) cDNA.
XX
KW   Multidrug resistance; MDR-1; P-glycoprotein;
KW   transmembrane efflux pump; haematopoietic stem cell; transduction;
KW   bone marrow transplantation; chemotherapy; radiation therapy; cancer;
KW   gene therapy; gene replacement; genetic defect; thalassaemia;
KW   Gauchier's disease; sickle cell anaemia; leukaemia; ex vivo expansion;
KW   cytokine; wild-type; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3843
FT                   /*tag=  a
FT                   /product= "Human wild-type MDR-1 protein"
FT   mutation        replace (553..555, GTT)
FT                   /*tag=  b
FT                   /note= "cDNA sequence of G185V human mutant MDR-1 given
FT                          in AAZ49333"
XX
PN   WO9961589-A2.
XX
PD   02-DEC-1999.
XX
PF   27-MAY-1999;   99WO-US11825.
XX
PR   28-MAY-1998;   98US-0086988.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI   Sorrentino B,  Bunting K;
XX
DR   WPI; 2000-072615/06.
DR   P-PSDB; AAY58186.
XX
PT   Ex vivo expansion of hematopoietic stem cells transduced with a
PT   sequence encoding human multidrug resistance-1, used for bone marrow
PT   transplantation -
XX
PS   Claim 10; Page 68-70; 113pp; English.
XX
CC   This sequence represents cDNA encoding human wild-type
CC   multidrug resistance protein MDR-1. MDR-1 is a transmembrane
CC   efflux pump, responsible for the export of drugs from cells,
CC   particularly cancer cells. Wild-type MDR-1 shows increased
CC   resistance to etoposide and decreased resistance to vinca
CC   alkaloids compared with a mutant form (AAY58187) where the Gly at
CC   position 185 is replaced by Val. The invention relates to transducing
CC   haematopoietic stem cells with nucleic acid encoding an MDR protein
CC   and culturing the modified cells. The modified haematopoietic stem
CC   cells are useful in bone marrow transplantation (to reconstitute
CC   haematopoietic systems in patients who have undergone chemotherapy or
CC   radiation therapy) and in ex vivo gene therapy of genetic defects in
CC   cells derived from haematopoietic stem cells, e.g., thalassaemia,
CC   Gauchier's disease, sickle cell anaemia or leukaemia. The modified
CC   cells can also be used to identify factors involved in regulating
CC   proliferation and differentiation in haematopoietic stem cells.
CC   Haematopoietic stem cells that express MDR-1 will be protected against
CC   chemotherapeutic agents, so can be engrafted while the patient is
CC   undergoing chemotherapy. Expansion of (rare) haematopoietic stem cells
CC   provides sufficient cells to permit standard biochemical analysis.
CC   Overexpression of MDR-1 allows cytokine-driven expansion of
CC   haematopoietic stem cells by at least 10-fold compared with a maximum
CC   of 4-fold in known procedures.
XX
SQ   Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;
```

```
Alignment Scores:
Pred. No.:               2.18e-154      Length:         3860
Score:                   1579.00        Matches:        306
Percent Similarity:      77.04%         Conservative:   100
Best Local Similarity:   58.06%         Mismatches:     99
Query Match:             60.92%         Indels:         22
DB:                      21             Gaps:           5

US-09-873-409-8 (1-514) x AAZ49332 (1-3860)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           ||:::::|||  ||||||::::::::::::||||||  |||||||||||   |||
Db     151 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210

Qy      21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
           |||  |||||||||:::|||                           :::|||:::|||
Db     211 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 270

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                       ::::::      |||      ||||||
Db     271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                 |||        ||||||  |||||||||        |||:::      |||
Db     331 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 390

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
           ||||||:::|||  |||   :::   |||   ||||||      :::|||||||||||
Db     391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
           |||||||||:::::::::     |||:::|||||||||||||||     |||:::||||||
Db     451 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 510

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
           ||||||||||||:::|||   |||:::   ||||||::::::||||||||||||||||||
Db     511 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGGTATTGGTGACAAAATT 570

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
              :::   ||||||:::|||:::||||||      |||      ||||||      :::
Db     571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 630

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
           ||||||||||||||||||||||   |||:::   ||||||::::::      |||||||||
Db     631 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
              ::::::::::::   |||   |||      |||      ||||||:::||||||
Db     691 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
           ||||||||||||||||||||||||::::::|||||||||||||||||||||      |||
Db     751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810

Qy     201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
           :::|||||||||:::||||||   :::||||||::::::||||||      |||||||||
Db     811 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 870

Qy     221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
           :::   |||   :::   :::|||:::||||||      :::      :::      :::|||
Db     871 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 930

Qy     241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
           |||||||||||||||||||||:::|||:::|||:::||||||       |||:::|||
Db     931 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 984

Qy     261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
           |||   ||||||   ||||||||||||||:::      ::::::   :::|||   |||
Db     985 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1044

Qy     281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
              |||      |||   ||||||   ||||||||||||:::   ||||||::::::
Db    1045 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104
```

```
Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db   1105 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1164

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db   1165 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1224

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   1225 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1284

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   1285 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1344

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   1345 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1404

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   1405 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1464

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1465 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1524

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1525 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1584

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   1585 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1644

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   1645 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1704

Qy    501 ValGlnAlaAlaLeuGluLys 507
Db   1705 GTTCAGGTGGCTCTGGATAAG 1725
```

```
RESULT 10
ABA94365
ID   ABA94365 standard; DNA; 3860 BP.
XX
AC   ABA94365;
XX
DT   26-MAR-2002  (first entry)
XX
DE   Human BCRP DNA related seq Id No. 1.
XX
KW   Stem cell; ATP transport protein; ATP-binding cassette; antiparkinsonian;
KW   hepatotropic; neurodegenerative; cytostatic; antianemic; muscular; BCRP;
KW   cardiant; gene therapy; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             1..3843
FT                   /*tag=  a
XX
PN   WO200192877-A2.
XX
PD   06-DEC-2001.
XX
PF   30-MAY-2001; 2001WO-US17459.
XX
PR   31-MAY-2000; 2000US-0584586.
PR   29-MAY-2001; 2001US-0866866.
XX
PA   (SJUD-) ST JUDE CHILDREN'S RES HOSPITAL.
XX
PI   Sorrentino B,  Schuetz J;
XX
DR   WPI; 2002-114368/15.
DR   P-PSDB; ABB07266.
XX
PT   Identifying a stem cell, for treating e.g., muscular dystrophy,
PT   myocardial infarction, Parkinson's disease, or neurodegenerative
PT   disorders, comprises detecting the expression of an ATP transport
PT   protein (BCRP) by a cell  -
XX
PS   Disclosure; Page 53-55; 87pp; English.
XX
CC   The invention provides a method of identifying and/or isolating a stem
CC   cell that involves detecting the expression of an ATP transport protein
CC   containing a conserved ATP-binding cassette (BCRP) by a cell in a sample
CC   comprising stem cells. The isolated stem cells may be used in the
CC   treatment of diseases such as muscular dystrophy, degenerative liver
CC   disorder, myocardial infarction, Parkinson's disease, degenerative
CC   disorders of the brain, and for tissue regeneration or replacement.
CC   Haematopoietic cells can be used in bone marrow transplants (e.g., for
CC   treatment of leukemia) and for ex vivo gene therapy for treating blood
CC   diseases such as sickle cell anemia and thalassemia. The stem cells can
CC   also be used as cell targets in gene therapy protocols. The present
CC   sequence represents a sequence related to the BCRP for which no relevant
CC   information has been provided in the specification.
XX
SQ   Sequence 3860 BP; 1135 A; 746 C; 957 G; 1022 T; 0 other;

Alignment Scores:
Pred. No.:              2.18e-154     Length:        3860
Score:                  1579.00       Matches:       306
Percent Similarity:     77.04%        Conservative:  100
Best Local Similarity:  58.06%        Mismatches:    99
Query Match:            60.92%        Indels:        22
DB:                     24            Gaps:          5

US-09-873-409-8 (1-514) x ABA94365 (1-3860)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    151 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 210

Qy     21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db    211 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 270

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db    271 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 330

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    331 ATGACCAGATATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 390

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    391 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 450

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    451 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 510

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    511 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGGTATTGGTGACAAAATT 570

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    571 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 630
```

```
Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    631 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 690

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    691 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 750

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    751 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 810

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db    811 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 870

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db    871 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 930

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db    931 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 984

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db    985 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1044

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db   1045 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1104

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db   1105 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1164

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db   1165 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1224

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   1225 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1284

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   1285 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1344

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   1345 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1404

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   1405 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1464

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1465 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1524

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1525 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1584

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   1585 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1644

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   1645 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1704

Qy    501 ValGlnAlaAlaLeuGluLys 507
Db   1705 GTTCAGGTGGCTCTGGATAAG 1725
```

RESULT 11
AAZ90198
ID AAZ90198 standard; DNA; 4233 BP.
XX
AC AAZ90198;
XX
DT 19-MAY-2000 (first entry)
XX
DE Rat mdr1b2 (multispecific drug transporter) nucleotide sequence.
XX
KW Rat; mdr1b2; multispecific drug transporter; cancer; inflammation;
KW cardiovascular disease; central nervous system disorder;
KW autoimmune disease; kidney disease; drug formulation; ss.
XX
OS Rattus sp.
XX
PN US6025160-A.
XX
PD 15-FEB-2000.
XX
PF 22-JUL-1998; 98US-0120513.
XX
PR 22-JUL-1998; 98US-0120513.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Brun KA, Ellens H, Yue L, Chenery RJ, Feild JA;
XX
DR WPI; 2000-181810/16.
DR P-PSDB; AAY78879.
XX
PT Isolated polynucleotide encoding a rat mdr1b2 drug transporter
PT polypeptide, useful for treatment of e.g. cancer, autoimmune disease,
PT central nervous system disorders -
XX
PS Claim 7; Column 13-17; 17pp; English.
XX
CC This sequence represents a polynucleotide encoding a 1275 amino acid
CC rat mdr1b2 multispecific drug transporter polypeptide. The mdr1b2
CC nucleotide sequence has cytostatic, antiinflammatory, cardiant,
CC neuroprotective, immunosuppressive and nephrotropic activity.
CC Understanding the functioning of the mdr1b2 polynucleotide and protein in
CC transgenic animal models is useful for treating and preventing diseases
CC such as cancer, inflammation, cardiovascular disease, central nervous
CC system disorders, autoimmune disease, and kidney disease. The use of the
CC protein in cell based, membrane based, or binding assays may enhance drug
CC formulation, selection of formulation excipients and compound design.
XX
SQ Sequence 4233 BP; 1210 A; 920 C; 1048 G; 1055 T; 0 other;

```
Alignment Scores:
Pred. No.:               5.12e-154     Length:          4233
Score:                   1576.00       Matches:         308
Percent Similarity:      75.90%        Conservative:    92
Best Local Similarity:   58.44%        Mismatches:      105
Query Match:             60.80%        Indels:          22
DB:                      21            Gaps:            4
```

US-09-873-409-8 (1-514) x AAZ90198 (1-4233)

```
Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    170 ATGGCTCTGGGAACTCTCGCTGCTATCATCCACGGAACCCTGCTTCCCCTCCTGATGCTG 229

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
Db    230 GTGTTCGGATACATGACAGATAGTTTTACCCAAGCAGAGACCCGCATTCTGCCGAGCGTT 289

Qy     31 ------SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu--------- 45
```

```
Db    290 ACTAATCAAAGTGAAATCAACAGTACACAGACCGTCAGCGACAGCAGTCTGGAGGAGGAC 349

Qy     46 ------------ThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    350 ATGGCCATGTACGCCTACTATTACACGGGCATTGGTGCCGGTGTGCTCATCGTTGCCTAC 409

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    410 ATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAAG 469

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    470 TTTTTCCATGCCATCATGAATCAGGAGATAGGCTGGTTTGACGTGAATGACGCTGGGGAG 529

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    530 CTCAACACCCGGCTCACAGATGACGTCTCCAAAATTAATGACGGAATTGGTGACAAACTT 589

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    590 GGAATGTTCTTTCAGTCCATAACGACATTTTCAGCCGGTTTTATAATAGGATTTATAAGT 649

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    650 GGTTGGAAGCTAACCCTTGTAATTTTGGCCGTCAGCCCTCTTATTGGGTTGTCATCTGCC 709

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    710 ATGTGGGCAAAGGTACTGACTTCATTTACTAATAAGGAACTCCAGGCTTATGCGAAAGCT 769

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    770 GGAGCAGTTGCCGAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCGTTTGGAGGACAA 829

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db    830 AAGAAGGAACTTGAAAGGTACAATAAAAATTTAGAAGAAGCTAAAAGAGTTGGCATAAAG 889

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db    890 AAAGCCATCACGGCCAACATTTCCATAGGTATTGCCTACCTGTTGGTCTATGCGTCTTAT 949

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db    950 GCACTGGCATTCTGGTATGGGACCTCCTTGGTCCTCTCAAATGAA------TATTCTATT 1003

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db   1004 GGACAAGTGCTTACCGTCTTCTTCTCTATTTTATTGGGGACTTTCAGTATTGGACATTTA 1063

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db   1064 GCCCCAAACATAGAAGCCTTTGCAAATGCAAGAGGGGCAGCCTATGAAATCTTCAAGATA 1123

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db   1124 ATTGATAATGAGCCAAGCATTGACAGCTTCTCAACCAAGGGACACAAACCAGACAGTATA 1183

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db   1184 ATGGGAAATTTGGAATTTAAAAATGTTTACTTCAACTACCCATCACGAAGTGAAGTTAAG 1243

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   1244 ATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTGGTTGGCAAC 1303

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   1304 AGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCATAGAG 1363

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   1364 GGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAA 1423

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   1424 ATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCCACCACGATTGCCGAAAACATT 1483

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1484 CGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAAT 1543

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1544 GCCTATGACTTCATCATGAAACTGCCCCACAAATTTAACACCCTGGTTGGTGAGAGAGGG 1603

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   1604 GCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAAC 1663

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   1664 CCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTG 1723

Qy    501 ValGlnAlaAlaLeuGluLys 507
Db   1724 GTTCAGGCCGCTCTGGATAAG 1744
```

RESULT 12
AAF27498
ID   AAF27498 standard; cDNA; 4233 BP.
XX
AC   AAF27498;
XX
DT   25-APR-2001 (first entry)
XX
DE   Rat mdr1b2 multidrug resistance cDNA, SEQ ID NO:1.
XX
KW   Rat; multidrug resistance; mdr1b2; multi-drug transporter family;
KW   drug screening; pharmacokinetic analysis; oral absorption;
KW   formulation design; bioavailability; transgenic animal; knockout animal;
KW   inflammation; cardiovascular disease; central nervous system disorder;
KW   cancer; autoimmune disorder; kidney disease; ss.
XX
OS   Rattus sp.
XX
PN   US6169166-B1.
XX
PD   02-JAN-2001.
XX
PF   29-NOV-1999;   99US-0450105.
XX
PR   22-JUL-1998;   98US-0120513.
XX
PA   (SMIK) SMITHKLINE BEECHAM CORP.
PA   (SMIK) SMITHKLINE BEECHAM PLC.
XX
PI   Chenery RJ, Ellens H, Brun KA, Yue L, Feild JA;
XX
DR   WPI; 2001-158280/16.
DR   P-PSDB; AAB60409.
XX
PT   New rat multidrug resistance protein (mdr1b2), useful for screening or
PT   identifying compounds that are (ant)agonists/inhibitors of the
PT   ratmdr1b2, as well as compounds with optimal development
PT   characteristics -
XX
PS   Disclosure; Column 13-14; 13pp; English.
XX
CC   The invention relates to a novel rat multidrug resistance (mdr)
CC   protein, mdr1b2 (AAB60409). Multi-specific drug transporter family
CC   proteins are present in cell which have a barrier function, such as
CC   intestinal epithelial cells, brain microvessel endothelial cells,
CC   kidney epithelial cells, and liver hepatocytes, and are also expressed
CC   by certain cancer cells. The rat mdr1b2 protein is useful for
CC   screening or identifying compounds that are agonists or antagonists

```
CC   of mdr1b2 activity. It may also be used to establish assays to
CC   predict oral absorption and pharmacokinetics of drugs in humans, and
CC   thus enhance the design of formulations through the identification of
CC   compounds with optimal development characteristics (i.e., high oral
CC   bioavailability, UID (once a day) dosing, reduced drug interactions,
CC   reduced variability, and reduced food effects), specifically to
CC   avoid interactions with human mdr-1. Transgenic and knockout animals
CC   created using DNA encoding the rat mdr1b2 may be used to gain an
CC   insight into treating and preventing human diseases such as cancer,
CC   inflammation, cardiovascular disease, central nervous system disorders,
CC   autoimmune disorders and kidney disease. The present sequence represents
CC   cDNA encoding rat mdr1b2.
XX
SQ   Sequence 4233 BP; 1210 A; 920 C; 1048 G; 1055 T; 0 other;

Alignment Scores:
Pred. No.:              5.12e-154    Length:          4233
Score:                  1576.00      Matches:         308
Percent Similarity:     75.90%       Conservative:    92
Best Local Similarity:  58.44%       Mismatches:      105
Query Match:            60.80%       Indels:          22
DB:                     22           Gaps:            4

US-09-873-409-8 (1-514) x AAF27498 (1-4233)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    170 ATGGCTCTGGGAACTCTCGCTGCTATCATCCACGGAACCCTGCTTCCCCTCCTGATGCTG 229

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
Db    230 GTGTTCGGATACATGACAGATAGTTTACCCAAGCAGAGACCCGCATTCTGCCGAGCGTT 289

Qy     31 ------SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu--------- 45
Db    290 ACTAATCAAAGTGAAATCAACAGTACACAGACCGTCAGCGACAGCAGTCTGGAGGAGGAC 349

Qy     46 ------------ThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    350 ATGGCCATGTACGCCTACTATTACACGGGCATTGGTGCCGGTGTGCTCATCGTTGCCTAC 409

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    410 ATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAAG 469

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    470 TTTTTCCATGCCATCATGAATCAGGAGATAGGCTGGTTTGACGTGAATGACGCTGGGGAG 529

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    530 CTCAACACCCGGCTCACAGATGACGTCTCCAAAATTAATGACGGAATTGGTGACAAACTT 589

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    590 GGAATGTTCTTTCAGTCCATAACGACATTTTCAGCCGGTTTTATAATAGGATTTATAAGT 649

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    650 GGTTGGAAGCTAACCCTTGTAATTTTGGCCGTCAGCCCTCTTATTGGGTTGTCATCTGCC 709

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    710 ATGTGGGCAAAGGTACTGACTTCATTTACTAATAAGGAACTCCAGGCTTATGCGAAAGCT 769

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    770 GGAGCAGTTGCCGAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCGTTTGGAGGACAA 829

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db    830 AAGAAGGAACTTGAAAGGTACAATAAAAATTTAGAAGAAGCTAAAAGAGTTGGCATAAAG 889

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db    890 AAAGCCATCACGGCCAACATTTCCATAGGTATTGCCTACCTGTTGGTCTATGCGTCTTAT 949

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db    950 GCACTGGCATTCTGGTATGGGACCTCCTTGGTCCTCTCAAATGAA------TATTCTATT 1003

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db   1004 GGACAAGTGCTTACCGTCTTCTTCTCTATTTTATTGGGGACTTTCAGTATTGGACATTTA 1063

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db   1064 GCCCCAAACATAGAAGCCTTTGCAAATGCAAGAGGGGCAGCCTATGAAATCTTCAAGATA 1123

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db   1124 ATTGATAATGAGCCAAGCATTGACAGCTTCTCAACCAAGGGACACAAACCAGACAGTATA 1183

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db   1184 ATGGGAAATTTGGAATTTAAAAATGTTTACTTCAACTACCCATCACGAAGTGAAGTTAAG 1243

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   1244 ATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTGGTTGGCAAC 1303

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   1304 AGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCATAGAG 1363

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   1364 GGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAA 1423

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   1424 ATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCCACCACGATTGCCGAAAACATT 1483

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1484 CGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAAT 1543

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1544 GCCTATGACTTCATCATGAAACTGCCCCACAAATTTAACACCCTGGTTGGTGAGAGAGGG 1603

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   1604 GCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAAC 1663

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   1664 CCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTG 1723

Qy    501 ValGlnAlaAlaLeuGluLys 507
Db   1724 GTTCAGGCCGCTCTGGATAAG 1744

RESULT 13
AAD03488
ID   AAD03488 standard; cDNA; 4279 BP.
XX
AC   AAD03488;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) cDNA #1.
XX
KW   Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   drug bioavailability; transgenic animal; genetic model; ss.
XX
```

```
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             17..3862
FT                   /*tag= a
FT                   /product= "Dog P-glycoprotein (PGP) #1"
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00303.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 3; Page 58-63; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) cDNA. This
CC   sequence is also referred as Genotype C cDNA. The
CC   PGP enzyme functions as an efflux pump exporting small molecules
CC   across the cell membrane. This enzyme is a member of the ABC
CC   transporter family.
XX
SQ   Sequence 4279 BP; 1294 A; 834 C; 1008 G; 1143 T; 0 other;

Alignment Scores:
Pred. No.:              5.2e-154      Length:         4279
Score:                  1576.00       Matches:        302
Percent Similarity:     75.61%        Conservative:   98
Best Local Similarity:  57.09%        Mismatches:     105
Query Match:            60.80%        Indels:         24
DB:                     22            Gaps:           4

US-09-873-409-8 (1-514) x AAD03488 (1-4279)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          |||:::::|||   :::|||::::::::::||||||  ||||||||||||   |||
Db    167 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 226

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
          |||   |||   |||:::|||:::      :::      :::
Db    227 GTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 286

Qy     35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                               :::        ||||||     |||
Db    287 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 346

Qy     45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                            ||||||   |||||||||          |||:::
Db    347 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 406

Qy     60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
             |||||||||:::|||   |||   :::   |||   ||||||     :::||||||
Db    407 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 466

Qy     80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
          ||||||||||||||||:::::::::   |||:::||||||||||||||||        |||:::
Db    467 AAACAATTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 526

Qy    100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
          ||||||||||||||||||:::|||     |||:::     ||||||:::::::||||||||||||
Db    527 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 586

Qy    119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
          ||||||   :::   |||   :::::::::||||||         |||      ||||||
Db    587 AAAATTGGAATGTTCTTTCACTCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 646

Qy    139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
             :::||||||||||||||||||||||||     |||:::     ||||||::::::       |||
Db    647 ACACGTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 706

Qy    159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
          ||||||      ::::::::::::   |||   |||   |||||||||   ||||||:::
Db    707 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 766

Qy    179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    767 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 826

Qy    199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
             |||:::|||||||||:::||||||   :::||||||::::::||||||        |||
Db    827 GGACAAAAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAGGAATTGGG 886

Qy    219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238
          ||||||:::   |||   :::   :::|||:::||||||   :::      :::
Db    887 ATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCA 946

Qy    239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
          :::|||   ||||||||||||||||||||||||:::|||:::   |||      |||
Db    947 TCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TAT 1000

Qy    259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
          :::||||||   ||||||   ||||||||||||||:::      ::::::     ||||||
Db   1001 TCTATTGGACAAGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGA 1060

Qy    279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
             |||   |||      |||   ||||||   ||||||||||||:::      ||||||
Db   1061 CAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTC 1120

Qy    299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
          ::::::||||||   ||||||||||||||||||::::::|||      :::|||:::||||||:::
Db   1121 AAGATAATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGAT 1180

Qy    319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
          :::|||:::|||   :::||||||||||||||||     |||:::||||||||||||
Db   1181 AATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAA 1240

Qy    339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
          :::||||||||||||||||||||||||:::::::::||||||:::||||||||||||||||
Db   1241 GTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTT 1300

Qy    359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
          |||      :::|||   ||||||||||||      ||||||||||||||||||||
Db   1301 GGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCC 1360

Qy    379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
             ||||||   :::   :::|||      ||||||||||      :::||||||||||||
Db   1361 ACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTT 1420

Qy    399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
          |||:::      ||||||||||||||||||||||||||||      ||||||||:::
Db   1421 CGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAA 1480
```

```
Qy     419 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
           |||||:::||||||||:::::||||||    :::|||:::|||:::|||    :::|||
Db    1481 AACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAA 1540

Qy     439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
           ||||||||||||||||||||||:::   ||||||||||:::||||||||||||||
Db    1541 GCCAATGCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAG 1600

Qy     459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
           :::||||||||:::||||||||||||||||||||||||||||||||||||||||||||||
Db    1601 AGAGGGGCCCAGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTT 1660

Qy     479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
           ||||||||||||||||:::||||||||||||||||||||||||:::||||||:::
Db    1661 CGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAA 1720

Qy     499 SerAlaValGlnAlaAlaLeuGluLys 507
           :::   ||||||   ||||||:::|||
Db    1721 GCAGTGGTTCAGGTGGCCCTGGATAAG 1747

RESULT 14
AAD03489
ID   AAD03489 standard; cDNA; 4317 BP.
XX
AC   AAD03489;
XX
DT   13-JUN-2001  (first entry)
XX
DE   Dog P-glycoprotein (PGP) cDNA #2.
XX
KW   Dog; P-glycoprotein; PGP; multidrug transporter; MDR1;
KW   drug bioavailability; transgenic animal; genetic model; ss.
XX
OS   Canis familiaris.
XX
FH   Key             Location/Qualifiers
FT   CDS             70..3912
FT                   /*tag= a
FT                   /product= "Dog P-glycoprotein (PGP) #2"
XX
PN   WO200123540-A2.
XX
PD   05-APR-2001.
XX
PF   28-SEP-2000; 2000WO-US26767.
XX
PR   28-SEP-1999;   99US-0156510.
XX
PA   (GENT-) GENTEST CORP.
XX
PI   Stocker PJ,  Steimel-crespi DT,  Crespi CL,  Reif TC,  Patten CJ;
XX
DR   WPI; 2001-235373/24.
DR   P-PSDB; AAE00304.
XX
PT   New dog P-glycoproteins (PGP) and their encoding nucleic acids, useful
PT   for determining the bioavailability of drugs and for screening for dog
PT   PGP inhibitors -
XX
PS   Claim 1; Page 66-72; 111pp; English.
XX
CC   The invention relates to dog P-glycoprotein (PGP) also referred
CC   as multidrug transporter (MDR1) and nucleic acids encoding them.
CC   The invention also includes fragments and biologically functional
CC   variants of dog P-glycoprotein. PGP and their nucleic acids are
CC   useful for determining the bioavailability of drugs and for
CC   screening PGP inhibitors. They are useful for the diagnosis and
CC   treatment of conditions characterised by PGP activity, by
CC   reducing or increasing PGP activity in a cell. PGP nucleic acids
CC   are used as oligonucleotide probes. Complements of PGP nucleic
CC   acids are useful as antisense oligonucleotides, to induce a PGP
CC   'knockout' phenotype. They are used to prepare a non-human
CC   transgenic animal, which are valuable as genetic models for
CC   human diseases.
CC   The present sequence is dog P-glycoprotein (PGP) cDNA. The
CC   PGP enzyme functions as an efflux pump exporting small molecules
CC   across the cell membrane. This enzyme is a member of the ABC
CC   transporter family.
XX
SQ   Sequence 4317 BP; 1293 A; 844 C; 1019 G; 1161 T; 0 other;

Alignment Scores:
Pred. No.:              5.27e-154     Length:         4317
Score:                  1576.00       Matches:        302
Percent Similarity:     75.61%        Conservative:   98
Best Local Similarity:  57.09%        Mismatches:     105
Query Match:            60.80%        Indels:         24
DB:                     22            Gaps:           4

US-09-873-409-8 (1-514) x AAD03489 (1-4317)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||::::::|||   :::|||::::::::::::||||||   ||||||||||||   |||
Db     217 ATGTTGGTGGGGACAATGGCTGCCATCATCCATGGAGCTGCACTCCCTCTCATGATGCTG 276

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu------------------ 34
           |||   |||   |||:::|||:::         :::      :::
Db     277 GTTTTTGGAAACATGACAGATAGCTTTGCAAATGCAGGAATTTCAAGAAACAAAACTTTT 336

Qy      35 ---------------------ValGlnThrAsnThrTyrSerPhePheArg--------- 44
                                :::      ||||||      |||
Db     337 CCAGTTATAATTAATGAAAGTATTACGAACAATACACAACATTTCATCAACCATCTGGAG 396

Qy      45 ---------------LeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePhe 59
                             ||||||   |||||||||            |||:::
Db     397 GAGGAAATGACCACGTATGCCTATTATTACAGTGGGATCGGTGCTGGCGTGCTGGTGGCT 456

Qy      60 GlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArg 79
              |||||||||:::|||   |||   :::   |||   ||||||      :::||||||
Db     457 GCTTACATCCAGGTTTCATTCTGGTGCCTGGCAGCAGGAAGACAGATACTCAAAATTAGA 516

Qy      80 LysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIle 99
           ||||||||||||||||||:::::::::   |||:::|||||||||||||||      |||:::
Db     517 AAACAATTTTTTCATGCTATCATGCGACAGGAGATTGGCTGGTTTGACGTGCATGACGTT 576

Qy     100 GlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAsp 118
           ||||||||||||||||||:::|||   |||:::   ||||||::::::||||||||||||
Db     577 GGGGAGCTTAACACCCGGCTCACAGACGATGTCTCCAAAATCAATGAAGGAATTGGCGAC 636

Qy     119 LysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu 138
           |||:::   :::   ||||||:::::::::||||||      |||      ||||||
Db     637 AAAGTTGGAATGTTCTTTCAATCAATAGCAACATTTTTCACCGGTTTTATAGTGGGGTTT 696

Qy     139 ValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSer 158
                    ||||||||||||||||||||   |||:::   ||||||::::::      |||
Db     697 ACACCTGGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGCCCTGTTCTTGGACTTTCA 756

Qy     159 AlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSer 178
           ||||||      ::::::::::::   |||   |||   |||||||||   ||||||:::
Db     757 GCCGCCATCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTGGCCTATGCA 816

Qy     179 LysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArg 198
           ||||||||||||||||||||||||||||||||||::::::||||||||||||||||||||
Db     817 AAAGCTGGAGCAGTAGCTGAAGAAGTCTTAGCAGCAATCAGAACTGTGATTGCCTTTGGA 876

Qy     199 AlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGly 218
              |||:::|||||||||:::||||||   :::||||||::::::||||||      |||
Db     877 GGACAAAAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGG 936

Qy     219 IleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGly 238
           ||||||:::   |||   :::   :::|||:::||||||   :::      :::
Db     937 ATAAAGAAAGCTATCACGGCCAACATTTCTATTGGTGCCGCTTTCTTATTGATCTATGCA 996

Qy     239 ThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyr 258
           :::|||   ||||||||||||||||||||||||:::|||:::   |||      |||
```

```
Db     997 TCATATGCTCTGGCTTTCTGGTATGGGACCTCCTTGGTCCTCTCCAGTGAA------TAT 1050

Qy     259 ThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGly 278
           ||||||||  ||||||  ||||||||||||||:::     ::::::  ||||||
Db    1051 ACTATTGGACAGGTACTCACTGTCTTCTTTTCTGTATTAATTGGGGCTTTTAGTATTGGA 1110

Qy     279 AlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePhe 298
              |||   |||      |||   ||||||   ||||||||||||||:::    ||||||
Db    1111 CAGGCATCCCCAAGCATTGAAGCATTTGCAAACGCAAGAGGAGCAGCTTATGAAATCTTC 1170

Qy     299 GlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGlu 318
           ::::::||||||   ||||||||||||||||:::::::|||   :::|||:::||||||:::
Db    1171 AAGATAATTGACAATAAACCAAGCATTGACAGCTATTCGAAGAGTGGACATAAACCAGAT 1230

Qy     319 SerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSer 338
           :::|||:::|||   :::||||||||||||||   |||:::|||||||||||
Db    1231 AATATTAAGGGAAATTTGGAATTCAAAAATGTTCACTTCAGTTACCCTTCTCGAAAAGAA 1290

Qy     339 IleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuVal 358
           :::||||||||||||||||||||||||:::::::::||||||:::||||||||||||||
Db    1291 GTTAAGATCTTAAAGGGTCTCAACCTGAAGGTTCAGAGTGGGCAGACAGTGGCGCTGGTT 1350

Qy     359 GlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspPro 378
           |||   :::|||   ||||||||||||   |||||||||:::|||||||||||||||||
Db    1351 GGGAACAGTGGCTGCGGGAAGAGCACGACCGTGCAGCTGATGCAGAGGCTCTATGACCCC 1410

Qy     379 AspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyr 398
              ||||||   :::   :::|||         ||||||||   :::||||||||||||
Db    1411 ACAGATGGCATGGTCTGTATTGATGGACAGGACATTAGGACCATAAATGTAAGGCATCTT 1470

Qy     399 ArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsn 418
           |||:::     |||||||||||||||||||||||||||||   |||||||||:::
Db    1471 CGGGAAATTACTGGTGTGGTGAGTCAGGAGCCTGTGTTGTTTGCCACCACGATAGCTGAA 1530

Qy     419 AsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGlu 438
           ||||||:::|||||||||:::::::||||||   :::|||:::|||:::|||   :::|||
Db    1531 AACATTCGCTATGGCCGCGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTTAAGGAA 1590

Qy     439 AlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGlu 458
           ||||||||||||||||||||||||:::   ||||||||||||:::||||||||||||||||
Db    1591 GCCAATGCCTATGATTTTATCATGAAACTACCTAATAAATTTGACACTCTGGTTGGAGAG 1650

Qy     459 LysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuVal 478
           :::||||||::::::||||||||||||||||||||||||||||||||||||||||||||
Db    1651 AGAGGGGCCCGGCTGAGTGGTGGACAGAAACAGAGAATCGCCATTGCTCGGGCCCTGGTT 1710

Qy     479 ArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLys 498
           |||||||||||||||||:::||||||||||||||||||||||||||||:::||||||:::
Db    1711 CGCAACCCCAAGATTCTTCTGCTGGATGAGGCAACGTCAGCTCTGGACACTGAAAGTGAA 1770

Qy     499 SerAlaValGlnAlaAlaLeuGluLys 507
           :::   ||||||   ||||||:::|||
Db    1771 GCAGTGGTTCAGGTGGCCCTGGATAAG 1797
```

RESULT 15
AAV32645

```
ID   AAV32645 standard; cDNA; 4669 BP.
XX
AC   AAV32645;
XX
DT   23-SEP-1998  (first entry)
XX
DE   Human P glycoprotein (Pgp) cDNA.
XX
KW   Human P glycoprotein; Pgp; multi-drug resistance; cancer;
KW   UIC2 monoclonal antibody; mAB; cytotoxic; transmembrane efflux pump; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   5'UTR           1..424
FT                   /*tag=  a
FT   CDS             425..4267
FT                   /*tag=  b
FT                   /product= "Wild-type Pgp"
FT   3'UTR           4268..4669
FT                   /*tag=  c
XX
PN   WO9821325-A1.
XX
PD   22-MAY-1998.
XX
PF   17-NOV-1997;   97WO-US21214.
XX
PR   15-NOV-1996;   96US-0752447.
XX
PA   (INGE-) INGENEX INC.
PA   (UNII ) UNIV ILLINOIS FOUND.
XX
PI   Mechetner E,  Roninson IB;
XX
DR   WPI; 1998-297930/26.
DR   P-PSDB; AAW48997.
XX
PT   Immunological reagent specific for P-glyco:protein - useful for
PT   detecting multi-drug resistant cancer, isolating haematopoietic
PT   cells and selective cell killing
XX
PS   Claim 1; Fig 1A; 89pp; English.
XX
CC   The present sequence represents the human P glycoprotein (Pgp)
CC   cDNA which encodes the wild-type Pgp protein.  Pgp is a transmembrane
CC   efflux pump protein involved in multi-drug resistance of cancer cells.
CC   The invention provides methods for developing and using immunological
CC   reagents specific for certain mutant forms of Pgp and wild-type Pgp in
CC   a conformation associated with substrate binding or in the presence
CC   of ATP depleting agents.  An example of the immunological reagent
CC   is the UIC2 monoclonal antibody (mAB).  mAB UIC2 specifically binds
CC   to Pgp in a particular biochemical conformation and is capable of
CC   inhibiting drug efflux from Pgp-expressing cells.  The immunological
CC   reagents are claimed to be useful for detecting Pgp expression in
CC   mammalian cells, including low level expression, particularly in cancer
CC   cells to diagnose multi-drug resistance.  The invention claims that
CC   these immunological reagents are more specific than known reagents for
CC   detecting Pgp and they also eliminate the need for costly and laborious
CC   screening of Pgp inhibitors by cytotoxicity or dye exclusion methods.
XX
SQ   Sequence 4669 BP; 1393 A; 894 C; 1130 G; 1252 T; 0 other;

Alignment Scores:
Pred. No.:              5.9e-154      Length:         4669
Score:                  1576.00       Matches:        305
Percent Similarity:     77.04%        Conservative:   101
Best Local Similarity:  57.87%        Mismatches:     99
Query Match:            60.80%        Indels:         22
DB:                     19            Gaps:           5

US-09-873-409-8 (1-514) x AAV32645 (1-4669)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||::::::|||   ||||||::::::::::::||||||   |||||||||||   |||
Db     575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy      21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
           |||   |||||||||:::|||                           :::|||:::|||
Db     635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                       ::::::      |||         ||||||
Db     695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                 |||         ||||||   |||||||||         |||:::      |||
Db     755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
```

```
Db     815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db     875 TTTTTCATGCTATTATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db     935 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAAATT 994

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db     995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy     201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db    1235 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1294

Qy     221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db    1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1354

Qy     241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db    1355 GCTGTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1408

Qy     261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db    1409 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1468

Qy     281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db    1469 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528

Qy     301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db    1529 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1588

Qy     321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db    1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1648

Qy     341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db    1649 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1708

Qy     361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db    1709 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1768

Qy     381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db    1769 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1828

Qy     401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db    1829 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1888

Qy     421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440

Db    1889 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1948

Qy     441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db    1949 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2008

Qy     461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db    2009 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2068

Qy     481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db    2069 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2128

Qy     501 ValGlnAlaAlaLeuGluLys 507
Db    2129 GTTCAGGTGGCTCTGGATAAG 2149
```

Search completed: March 30, 2003, 03:22:28
Job time : 555.349 secs

```
OM protein - nucleic search, using frame_plus_p2n model

Run on:          March 30, 2003, 03:22:44 ; Search time 3535.38 Seconds
                                            (without alignments)
                                            2354.618 Million cell updates/sec

Title:           US-09-873-409-8
Perfect score:   2592
Sequence:        1 MILGILASLVNGACLPLMPL..........SESKSAVQAALEKDTPRYSF 514

Scoring table:   BLOSUM62
                 Xgapop 10.0 , Xgapext  0.5
                 Ygapop 10.0 , Ygapext  0.5
                 Fgapop  6.0 , Fgapext  7.0
                 Delop   6.0 , Delext   7.0

Searched:        16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters:      32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115420_19240/app_query.fasta_1.7544
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_30544_@runat_27032003_115420_19240 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :       EST:*
                 1:  em_estba:*
                 2:  em_esthum:*
                 3:  em_estin:*
                 4:  em_estmu:*
                 5:  em_estov:*
                 6:  em_estpl:*
                 7:  em_estro:*
                 8:  em_htc:*
                 9:  gb_est1:*
                 10:  gb_est2:*
                 11:  gb_htc:*
                 12:  gb_est3:*
                 13:  gb_est4:*
                 14:  gb_est5:*
                 15:  em_estfun:*
                 16:  em_estom:*
                 17:  gb_gss:*
                 18:  em_gss_hum:*
                 19:  em_gss_inv:*
                 20:  em_gss_pln:*
                 21:  em_gss_vrt:*
                 22:  em_gss_fun:*
                 23:  em_gss_mam:*
                 24:  em_gss_mus:*
                 25:  em_gss_other:*
                 26:  em_gss_pro:*
                 27:  em_gss_rod:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1092	42.1	669	13	BM013981	BM013981 603639619
	2	1068.5	41.2	1033	14	BM926413	BM926413 AGENCOURT
	3	1029	39.7	760	13	BM016204	BM016204 603642659
	4	926.5	35.7	929	14	BQ882401	BQ882401 AGENCOURT
	5	918	35.4	545	9	AL040762	AL040762 DKFZp434C
	6	856.5	33.0	948	13	BM471690	BM471690 AGENCOURT
	7	750	28.9	894	12	BF584668	BF584668 602098406
	8	744.5	28.7	872	11	AK020318	AK020318 Mus muscu
	9	710	27.4	2676	11	AK014319	AK014319 Mus muscu
	10	671.5	25.9	803	12	BG298756	BG298756 602396681
	11	659	25.4	943	9	AL520322	AL520322 AL520322
	12	654	25.2	405	9	AA243820	AA243820 zr67g06.r
	13	650	25.1	550	13	BJ075336	BJ075336 BJ075336
	14	637	24.6	780	12	BG587938	BG587938 EST489713
	15	625	24.1	559	12	BF692596	BF692596 602248949
c	16	621	24.0	606	14	BQ262340	BQ262340 fz85a08.x
	17	621	24.0	681	14	BU006064	BU006064 QGG9N16.y
c	18	604.5	23.3	932	17	AZ670821	AZ670821 ENTJN69TF
	19	602.5	23.2	913	17	BH155700	BH155700 ENTRO54TR
c	20	599	23.1	880	17	AZ687805	AZ687805 ENTLU52TF
	21	596.5	23.0	871	17	AZ682350	AZ682350 ENTKB16TF
	22	596.5	23.0	2564	11	AK020022	AK020022 Mus muscu
c	23	590.5	22.8	891	17	AZ682250	AZ682250 ENTKT68TR
c	24	585.5	22.6	853	17	AZ679807	AZ679807 ENTHI67TR
	25	584.5	22.6	1813	11	U66688	U66688 Homo sapien
	26	582.5	22.5	787	12	BG584063	BG584063 EST485823
	27	581	22.4	586	13	BJ336461	BJ336461 BJ336461
c	28	579	22.3	661	12	BG302675	BG302675 f136b04.x
c	29	578	22.3	575	10	AW342644	AW342644 fi81f06.x
	30	577.5	22.3	886	17	AZ540627	AZ540627 ENTEQ18TF
c	31	577	22.3	835	13	BJ349604	BJ349604 BJ349604
	32	576.5	22.2	947	17	AZ683753	AZ683753 ENTIL96TF
	33	576	22.2	547	13	BJ333166	BJ333166 BJ333166
c	34	574.5	22.2	885	17	AZ686798	AZ686798 ENTFZ28TR
	35	574	22.1	542	13	BJ339421	BJ339421 BJ339421
c	36	572.5	22.1	860	17	AZ543293	AZ543293 ENTDQ80TF
	37	572.5	22.1	886	17	BH139685	BH139685 ENTNA47TR
	38	571	22.0	715	14	BQ869512	BQ869512 QGD6G09.y
c	39	570	22.0	814	13	BJ352304	BJ352304 BJ352304
c	40	570	22.0	820	13	BJ356540	BJ356540 BJ356540
	41	570	22.0	1019	12	BG248052	BG248052 602359987
	42	569	22.0	1341	11	AY108285	AY108285 Zea mays
	43	565	21.8	657	14	BQ802058	BQ802058 WHE2821_F
c	44	564	21.8	636	13	BJ386538	BJ386538 BJ386538
c	45	564	21.8	750	13	BJ445883	BJ445883 BJ445883

ALIGNMENTS

```
RESULT 1
BM013981
LOCUS       BM013981                669 bp    mRNA    linear   EST 30-OCT-2001
DEFINITION  603639619F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5415556 5',
            mRNA sequence.
ACCESSION   BM013981
VERSION     BM013981.1  GI:16528335
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 669)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
```

```
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12060  row: f  column: 05
            High quality sequence stop: 664.
FEATURES             Location/Qualifiers
     source          1..669
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5415556"
                     /clone_lib="NIH_MGC_87"
                     /tissue_type="mammary adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.383 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      203 a    122 c    162 g    182 t
ORIGIN

Alignment Scores:
Pred. No.:              1.77e-119     Length:          669
Score:                  1092.00       Matches:         215
Percent Similarity:     99.54%        Conservative:    2
Best Local Similarity:  98.62%        Mismatches:      0
Query Match:            42.13%        Indels:          1
DB:                     13            Gaps:            0

US-09-873-409-8 (1-514) x BM013981 (1-669)

Qy      265 AlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPhe 284
            :::|||||||||:::|||||||||||||||||||||||||||||||||||||||||||||
Db       17 AGTGTTTTCTTTAATGTAATCCATAGCAGTTATTGCATTGGAGCAGCAGTCCCTCACTTT 76

Qy      285 GluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLys 304
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       77 GAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTTATTGATAAGAAA 136

Qy      305 ProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrVal 324
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      137 CCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTG 196

Qy      325 GluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGly 344
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      197 GAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGT 256

Qy      345 LeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGly 364
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      257 CTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGG 316

Qy      365 LysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMet 384
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      317 AAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATG 376

Qy      385 ValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyVal 404
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      377 GTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTG 436

Qy      405 ValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArg 424
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      437 GTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGA 496

Qy      425 AspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPhe 444
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      497 GATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTT 556

Qy      445 IleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSer 464
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      557 ATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATGAGT 616

Qy      465 GlyGlyGlnLysGlnArg-IleAlaIleAlaArgAlaLeuValArgAsnPro 481
            |||||||||||||||||| |||||||||||||||||||||||||||||||||
Db      617 GGAGGGCAGAAACAGAGGCATCGCAATTGCTCGTGCCTTAGTTCGAAACCCC 668

RESULT 2
BM926413
LOCUS       BM926413                1033 bp    mRNA    linear   EST 12-MAR-2002
DEFINITION  AGENCOURT_6600788 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5764845
            5', mRNA sequence.
ACCESSION   BM926413
VERSION     BM926413.1  GI:19376792
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1033)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12818  row: o  column: 22
            High quality sequence start: 3
            High quality sequence stop: 686.
FEATURES             Location/Qualifiers
     source          1..1033
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5764845"
                     /clone_lib="NIH_MGC_114"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: EcoRV (destroyed); RNA source anonymous pool of 6
                     male brains, age range 23-27 yo. Library is oligo-dT
                     primed and directionally cloned (EcoRV site is  destroyed
                     upon cloning).  Average insert size 1.5 kb, insert size
                     range 1-3 kb.  Library is normalized and enriched for
                     full-length  clones and was constructed by C. Gruber
                     (Invitrogen).  Research  Genetics tracking code 019. Note:
                     this is a NIH_MGC Library."
BASE COUNT      289 a    212 c    218 g    312 t      2 others
ORIGIN

Alignment Scores:
Pred. No.:              2.3e-116      Length:          1033
Score:                  1068.50       Matches:         229
Percent Similarity:     93.17%        Conservative:    3
Best Local Similarity:  91.97%        Mismatches:      17
Query Match:            41.22%        Indels:          4
DB:                     14            Gaps:            0

US-09-873-409-8 (1-514) x BM926413 (1-1033)

Qy      266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      180 GTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAA 239

Qy      286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      240 ACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTCATTGATAAGAAACCC 299
```

```
Qy     306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     300 AGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGAA 359

Qy     326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     360 TTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCTG 419

Qy     346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     420 AATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAG 479

Qy     366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     480 AGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGTG 539

Qy     386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     540 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTGGTT 599

Qy     406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     600 AGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGAGAT 659

Qy     426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     660 GATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATC 719

Qy     446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465
           ||||||||||||||||| |||||||||||||||||||||   ||||||   |||||||||
Db     720 ATGGAGTTTCCTAATAA-TTTAATACATTGGTAGGGGAANAAGGAGCTCCAATGAGTGGA 778

Qy     466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
              |||     ||||||  ||||||  ||||||||||||||||:::    |||
Db     779 AGCCAGAACCG-AGGATCCCAATTGCTCCTGCCTTAGTTCGAAACCCCAGGAT-CTGAAT 836

Qy     486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeu 505
           ||||||||||||||||||||||||||||||:::           |||   |||   |||
Db     837 TTAGATGAAGCTACGTCTGCCCTGGATTCAGAAACCAGTCAACTGGTCA--GCTGGCCTG 894

Qy     506 GluLysAspThrProArgTyrSerPhe 514
           ||||||:::|||||||||      |||
Db     895 GAGAAGGAAACCCCCCGGNAATTCTTT 921
```

RESULT 3
BM016204

```
LOCUS       BM016204                760 bp    mRNA    linear   EST 30-OCT-2001
DEFINITION  603642659F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5418615 5',
            mRNA sequence.
ACCESSION   BM016204
VERSION     BM016204.1  GI:16530558
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 760)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12068  row: e  column: 16
            High quality sequence stop: 740.
FEATURES             Location/Qualifiers
     source          1..760
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5418615"
                     /clone_lib="NIH_MGC_87"
                     /tissue_type="mammary adenocarcinoma, cell line"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: breast; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                     Average insert size 1.383 kb. Library enriched for
                     full-length clones and constructed by Life Technologies.
                     Note: this is a NIH_MGC Library."
BASE COUNT      229 a    150 c    187 g    193 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              7.18e-112      Length:        760
Score:                  1029.00        Matches:       209
Percent Similarity:     98.12%         Conservative:  0
Best Local Similarity:  98.12%         Mismatches:    2
Query Match:            39.70%         Indels:        2
DB:                     13             Gaps:          0

US-09-873-409-8 (1-514) x BM016204 (1-760)

Qy     303 LysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGly 322
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3 AAGAAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGA 62

Qy     323 ThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeu 342
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63 ACTGTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTG 122

Qy     343 LysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGly 362
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 AAAGGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGC 182

Qy     363 SerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPhe 382
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 AGTGGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTT 242

Qy     383 IleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle 402
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 ATCATGGTGGATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATN 302

Qy     403 GlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr 422
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 GGAGTGGTTAGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTAT 362

Qy     423 GlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyr 442
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     363 GGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTAT 422

Qy     443 AspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGln 462
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     423 GATTTTATCATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAA 482

Qy     463 MetSerGlyGlyGlnLysGlnArgIleAlaIle-AlaArgAlaLeuValArgAsnProLy 482
           |||||||||||||||||||||||||||||||| |||||||||||||||||||||||||||
Db     483 ATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGGCTCGTGCCTTAGTTCGAAACCCCAA 542

Qy     482 sIleLeuIleLeu-AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValG 502
           |||||||||     ||||||||||||||||||||||||||||||||||||||||||||||
Db     543 GATTCTGATTTTCAGATGAGGCTACGTCTGCCCTGGATTCAGAAAGCAAGTCAGCTGTTC 602

Qy     502 lnAlaAlaLeuGluLysAspThrProArgTyrSer 513
           ||||||||||||||||||||||||||||||||||
Db     603 AAGCTGCACTGGAGAAGGATACCCCCAGGTATTCC 637
```

RESULT 4
BQ882401

```
LOCUS       BQ882401                929 bp    mRNA    linear   EST 16-AUG-2002
DEFINITION  AGENCOURT_8627902 NIH_MGC_43 Homo sapiens cDNA clone IMAGE:6291782
```

```
             5', mRNA sequence.
ACCESSION    BQ882401
VERSION      BQ882401.1  GI:22274409
KEYWORDS     EST.
SOURCE       human.
  ORGANISM   Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE    1  (bases 1 to 929)
  AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
  TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL    Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
             Email: cgapbs-r@mail.nih.gov
             Tissue Procurement: ATCC
              cDNA Library Preparation: Rubin Laboratory
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Agencourt Bioscience Corporation
              Clone distribution: MGC clone distribution information can be
             found through the I.M.A.G.E. Consortium/LLNL at:
             http://image.llnl.gov
             Plate: LLCM2492  row: c  column: 15
             High quality sequence stop: 677.
FEATURES              Location/Qualifiers
     source           1..929
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:6291782"
                      /clone_lib="NIH_MGC_43"
                      /tissue_type="normal pigmented retinal epithelium"
                      /lab_host="DH10B (phage-resistant)"
                      /note="Organ: eye; Vector: pOTB7; Site_1: XhoI; Site_2:
                      EcoRI; cDNA made by oligo-dT priming.  Directionally
                      cloned into EcoRI/XhoI sites using the following 5'
                      adaptor: GGCACGAG(G).  Library constructed  by Ling Hong
                      in the laboratory of Gerald M. Rubin (University of
                      California, Berkeley) using ZAP-cDNA synthesis kit
                      (Stratagene) and  Superscript II RT (Life Technologies).
                      Note: this is a NIH_MGC Library. |"
BASE COUNT      281 a    176 c    237 g    234 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              1.71e-99      Length:         929
Score:                  926.50        Matches:        188
Percent Similarity:     79.30%        Conservative:   61
Best Local Similarity:  59.87%        Mismatches:     58
Query Match:            35.74%        Indels:         7
DB:                     14            Gaps:           2

US-09-873-409-8 (1-514) x BQ882401 (1-929)

Qy     176 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThr-ValIl 195
           ||||||:::|||||||||||||||||||||||||||||||::::::|||||||| |||||
Db       1 GCGTATGCAAAAGCTGGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGGTGAT 60

Qy     195 eAlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLy 215
            |||||||        |||:::||||||||||||     :::||||||::::::|||||
Db      61 TGCATTTGGAGGACAAAAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAA 120

Qy     215 sAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePh 235
            |      |||||||||:::  |||   :::   :::|||:::|||||   :::
Db     121 AAGAATTGGGATAAAGAAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCT 180

Qy     235 eMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGl 255
            :::       :::|||   ||||||||||||||||||||:::|||:::|||:::|||||
Db     181 GATCTATGCATCTTATGCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGA 240

Qy     255 uProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTy 275
            |      |||:::||||||   ||||||   |||||||||||||||:::       :::::
Db     241 A------TATTCTATTGGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTT 294

Qy     275 rCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPh 295
            :   :::|||   |||   |||        |||   ||||||    ||||||||||||||::
Db     295 TAGTGTTGGACAGGCATCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTA 354

Qy     295 eHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTy 315
           :    ||||||::::::||||||    |||||||||||||||::::::|||     :::|||::
Db     355 TGAAATCTTCAAGATAATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCA 414

Qy     315 rLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSe 335
           :||||||:::::|||:::|||     :::||||||:::||||||      |||:::|||||||||
Db     415 CAAACCAGATAATATTAAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATC 474

Qy     335 rArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrVa 355
           ||||      :::||||||||||||||||||||||||||||:::::::::||||||:::|||||
Db     475 TCGAAAAGAAGTTAAGATCTTGAAGGGTCTGAACCTGAAGGTGCAGAGTGGGCAGACGGT 534

Qy     355 lAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLe 375
           ||||||||||||      :::|||    ||||||||||||     |||||||||:::||||||||
Db     535 GGCCCTGGTTGGAAACAGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCT 594

Qy     375 uTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVa 395
           |||||||||     :::|||     :::   ||||||         ||||||||     :::|||||
Db     595 CTATGACCCCACAGAGGGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGT 654

Qy     395 lArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrTh 415
           ||||      |||:::   ||||||||||||||||||||||||||||||      |||||
Db     655 AAGGTTTCTACGGGAAATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCAC 714

Qy     415 rIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAl 435
           ||||:::     ||||||     |||||||||::::::||||||      :::|||:::|||:::||
Db     715 GATAGCTGAAAACATTCCCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGC 774

Qy     435 aAlaArgGluAlaAsnAla-TyrAspPheIleMetGluPheProAsnLysPheAsnThrL 455
           |     :::|||||||||||| ||||||||||||||:::     |||:::||||||:::||||
Db     775 TGTCCAGGAAGCCAATGCCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCC 834

Qy     455 euValGlyGluLysGlyAlaGlnMetSerGlyGly---GlnLysGlnArgIleAlaIleA 474
           ||||||||     :::        |||:::|||||||||     |||     :::          :::|
Db     835 TGGTTGGANAAAGAAGGGGCCAATTGAGTGGTGGGGCACAAACCAAAGGATCTCCCTTGG 894

Qy     474 laArg-AlaLeuVal-ArgAsnProLysIleLeu 484
           ||||| |||||||||     ||||||||||||||
Db     895 CACGTGGCCCTGGTTCCCCAACCCCAAGATCCTC 928

RESULT 5
AL040762
LOCUS       AL040762               545 bp    mRNA    linear   EST 29-FEB-2000
DEFINITION  DKFZp434C1815_r1 434 (synonym: htes3) Homo sapiens cDNA clone
            DKFZp434C1815 5', mRNA sequence.
ACCESSION   AL040762
VERSION     AL040762.1  GI:5409708
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 545)
  AUTHORS   Bloecker,H., Boecher,M., Brandt,P., Mewes,H.W., Gassenhuber,J. and
            Wiemann,S.
  TITLE     EST (Bloecker, et al.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Bloecker H
            MIPS
            Am Klopferspitz 18a D-82152 Martinsried, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by GBF (National Research Centre for Biotechnology Ltd.,
            Braunschweig/Germany) within the cDNA sequencing consortium of the
            German Genome Project.
            s1 sequence also available.
            This clone (DKFZp434C1815) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
```

```
FEATURES             Location/Qualifiers
     source          1. .545
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp434C1815"
                     /clone_lib="434 (synonym: htes3)"
                     /tissue_type="testis"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT      171 a     89 c    138 g    147 t
ORIGIN

Alignment Scores:
Pred. No.:                7.34e-99      Length:         545
Score:                    918.00        Matches:        181
Percent Similarity:       100.00%       Conservative:   0
Best Local Similarity:    100.00%       Mismatches:     0
Query Match:              35.42%        Indels:         0
DB:                       9             Gaps:           0

US-09-873-409-8 (1-514) x AL040762 (1-545)

Qy     306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       3 AGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGAA 62

Qy     326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      63 TTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCTG 122

Qy     346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     123 AATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAG 182

Qy     366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     183 AGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGTG 242

Qy     386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     243 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTGGTT 302

Qy     406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     303 AGTCAAGAGCCTGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGACGAGAT 362

Qy     426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     363 GATGTGACTGATGAAGAGATGGAGAGAGCAGCAAGGGAAGCAAATGCGTATGATTTTATC 422

Qy     446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     423 ATGGAGTTTCCTAATAAATTTAATACATTGGTAGGGGAAAAAGGAGCTCAAATGAGTGGA 482

Qy     466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     483 GGGCAGAAACAGAGGATCGCAATTGCTCGTGCCTTAGTTCGAAACCCCAAGATTCTGATT 542

Qy     486 Leu 486
           |||
Db     543 TTA 545

RESULT 6
BM471690
LOCUS       BM471690                 948 bp    mRNA    linear   EST 05-FEB-2002
DEFINITION  AGENCOURT_6465349 NIH_MGC_72 Homo sapiens cDNA clone IMAGE:5539117
            5', mRNA sequence.
ACCESSION   BM471690
VERSION     BM471690.1  GI:18520732
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 948)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM12233  row: b  column: 14
            High quality sequence stop: 569.
FEATURES             Location/Qualifiers
     source          1. .948
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:5539117"
                     /clone_lib="NIH_MGC_72"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 2 kb.   Library constructed by Life
                     Technologies."
BASE COUNT      285 a    187 c    194 g    282 t
ORIGIN

Alignment Scores:
Pred. No.:                3.98e-91      Length:         948
Score:                    856.50        Matches:        186
Percent Similarity:       89.67%        Conservative:   5
Best Local Similarity:    87.32%        Mismatches:     14
Query Match:              33.04%        Indels:         8
DB:                       13            Gaps:           1

US-09-873-409-8 (1-514) x BM471690 (1-948)

Qy     266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     205 GTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATTGGAGCAGCAGTCCCTCACTTTGAA 264

Qy     286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     265 ACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTTATTGATAAGAAACCC 324

Qy     306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     325 AGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACTGTGGAA 384

Qy     326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     385 TTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAAGGTCTG 444

Qy     346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     445 AATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGTGGGAAG 504

Qy     366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     505 AGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATCATGGTG 564

Qy     386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIle-GlyValVa 405
           ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Db     565 GATGAGAATGACATCAGAGCTTTAAATGTGCGGCATTATCGAGACCATATTGGAGTGGT 624

Qy     405 lSerGlnGluPro--ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyr-GlyAr 424
           ||||||||||||  ||||||||||||||||||||||||||||||||||||||| |||||
Db     625 TAGTCAAGAGCCCTGGTTTTGTTCGGGACCACCATCAGTAACAATATCAAGTATGGGACG 684
```

```
Qy     424 gAspAspValThrAspGluGlu-MetGluArgAlaAlaArgGluAla-AsnAlaTyrAsp 443
            ||||||  ||||||||:::  ||||||:::|||  ||||||       |||   :::
Db     685 AGATGATGGGACTGATGAACAAAATGGAAAAAGCAACAAGGGAAACCAAAGGCGAATGAA 744

Qy     444 PheIleMetGlu-PheProAsnLysPheAsnThrLeuVal---GlyGluLysGlyAlaGl 462
            ||||||||||| ||||||  ||||||||||||||         |||:::|||      ||
Db     745 TTTATCATGGAATTTCCCAAAAAAATTTAATACATTGGAAAGGGGAAAAAAGGACCTTCA 804

Qy     462 nMetSerGlyGlyGlnLysGlnArg 470
            |        |||   |||   :::
Db     805 AGGAATGGAAGGGCCAAAACCCAAG 829
```

```
RESULT 7
BF584668
LOCUS       BF584668                 894 bp    mRNA    linear   EST 12-DEC-2000
DEFINITION  602098406F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4218385 5',
            mRNA sequence.
ACCESSION   BF584668
VERSION     BF584668.1  GI:11658386
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 894)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Jeffrey E. Green, M.D.
             cDNA Library Preparation: Life Technologies, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLAM9798  row: 1  column: 02
            High quality sequence stop: 651.
FEATURES             Location/Qualifiers
     source          1..894
                     /organism="Mus musculus"
                     /strain="FVB/N"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:4218385"
                     /clone_lib="NCI_CGAP_Co24"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
                     Site_2: SalI; Cloned unidirectionally.  Primer: Oligo dT.
                     Average insert size 1.6 kb. Constructed by Life
                     Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT      263 a    203 c    242 g    185 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:               1.85e-78       Length:         894
Score:                   750.00         Matches:        156
Percent Similarity:      84.94%         Conservative:   47
Best Local Similarity:   65.27%         Mismatches:     36
Query Match:             28.94%         Indels:         4
DB:                      12             Gaps:           0

US-09-873-409-8 (1-514) x BF584668 (1-894)
```

```
Qy     266 ValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGlu 285
           ||||||||||||||:::      ::::::   :::|||   |||   |||:::   |||
Db       5 GTCTTCTTTTCCGTGTTAAT-GGAGCATTCAGTGTTGGACAGGCATCTCCAAATATTGAA 63

Qy     286 ThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysPro 305
              ||||||   ||||||||||||||:::   :::|||::::::||||||   ||||||
Db      64 GCCTTCGCCAATGCACGAGGAGCAGCTTATGAAGTCTTCAAAATAATTGATAATAAGCCC 123

Qy     306 SerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGlu 325
           |||||||:::||||||   :::|||:::||||||::::::|||:::|||   :::|||
Db     124 AGTATAGACAGCTTCTCAAAGAGTGGGCACAAACCAGACAACATACAAGGAAATCTGGAA 183

Qy     326 PheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeu 345
           |||||||:::   |||:::||||||||||||         ::::::||||||||||||||
Db     184 TTTAAGAATATTCACTTCAGTTACCCATCTCGAAAAGAAGTTCAGATCTTGAAGGGCCTC 243

Qy     346 AsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLys 365
           ||||||::::::|||||||||:::||||||||||||||||||   :::|||   ||||||
Db     244 AATCTGAAGGTGAAGAGCGGACAGACGGTGGCCCTGGTTGGCAACAGTGGCTGTGGAAAA 303

Qy     366 SerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetVal 385
           ||||||   |||||||||:::||||||||||||||||||   ||||||   :::   :::
Db     304 AGCACAACTGTCCAGCTGATGCAAAGGCTCTACGACCCCCTAGATGGCATGGTCAGTATC 363

Qy     386 AspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValVal 405
           |||      |||||||||   :::|||||||||:::   |||:::   ||||||||||||
Db     364 GACGGACAGGACATCAGAACCATCAATGTGAGGTATCTGAGGGAGATCATTGGTGTGGTG 423

Qy     406 SerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAsp 425
           ||||||||||||||||||||||   |||||||||||||||:::||||||||||||:::
Db     424 AGTCAGGAACCTGTGCTGTTTGCCACCACGATCGCCGAGAACATTCGCTATGGCCGAGAA 483

Qy     426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
           |||||||||   :::|||:::|||:::|||   :::||||||||||||||||||||||||
Db     484 GATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAATGCCTATGACTTCATC 543

Qy     446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465
           |||:::   |||::::::|||:::|||||||||||||||:::|||||||||:::||||||
Db     544 ATGAAACTGCCCCACCAATTTGACACCCTGGTTGGTGAGAGAGGGGCGCAGCTGAGTGGG 603

Qy     466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
           ||||||||||||||||||   ||||||||||||||||||||||||||||   |||
Db     604 GGACAGAAACAGAGAATCGC-ATTGCCCGGGCCCTGGTCCGCAATCCCAAGTGCCTTTC- 661

Qy     486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAla 504
           :::|||||||||||||||   ||||||:::||||||::::::   ||||||||||||
Db     662 GTGGACGAAGCCACCTCAGC-CTGGATACAGAAAGTGAAGCTGGGGTTCAGGCCGCA 717
```

```
RESULT 8
AK020318
LOCUS       AK020318                 872 bp    mRNA    linear   HTC 19-JAN-2002
DEFINITION  Mus musculus adult male epididymis cDNA, RIKEN full-length enriched
            library, clone:9230106F14:similar to MULTIDRUG RESISTANCE PROTEIN 3
            (P-GLYCOPROTEIN 3) (MDR1A), full insert sequence.
ACCESSION   AK020318
VERSION     AK020318.1  GI:12860872
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (strain:C57BL/6J) adult male epididymis cDNA to mRNA,
            clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:9230106F14.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
```

```
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409 (6821), 685-690 (2001)
  MEDLINE   21085660
   PUBMED   11217851
REFERENCE   5  (bases 1 to 872)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C.,
            Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,
            Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F.,
            Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,
            Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,
            Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
            Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,
            Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and
            Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Gemomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. cDNA went
            through one round of normalization to Rot = 10.0 and subtraction to
            Rot = 185.2. Second strand cDNA was prepared with the primer
            adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cleaved
            with BamHI and XhoI. Vector: a modified pBluescript KS(+) after
            bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'
            end: BamHI. Host: DH10B.
```

```
FEATURES             Location/Qualifiers
     source          1. .872
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:9230106F14"
                     /db_xref="MGD:MGI:1911565"
                     /db_xref="taxon:10090"
                     /clone="9230106F14"
                     /sex="male"
                     /tissue_type="epididymis"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="adult"
     CDS             86. .>872
                     /note="data source:SPTR, source key:P21447, evidence:ISS
                     putative
                     similar to MULTIDRUG RESISTANCE PROTEIN 3 (P-GLYCOPROTEIN
                     3) (MDR1A)"
                     /codon_start=1
                     /protein_id="BAB32070.1"
                     /db_xref="GI:12860873"
                     /db_xref="MGD:MGI:1924956"
                     /translation="MANSERTNGLQETNQRYGPLQEQVPKVGNQAVGPIEIFRFADNL
                     DIVLMTLGILASMINGATVPLMSLVLGEISDHLINGCLVQTNRTKYQNCSQTQEKLNE
                     DIIVLTLYYIGIGAAALIFGYVQISFWVITAARQTTRIRKQFFHSILAQDISWFDGSD
                     ICELNTRMTGDINKLCDGIGDKIPLMFQNISGFSIGLVISLIKSWKLSLVVLSTSPLI
                     MASSALCSRMIISLTSKELDAYSKAGAVAEEALSSIQTVTAFGA"
BASE COUNT      262 a    178 c    187 g    245 t
ORIGIN

Alignment Scores:
Pred. No.:              8.06e-78      Length:         872
Score:                  744.50        Matches:        151
Percent Similarity:     80.37%        Conservative:   21
Best Local Similarity:  70.56%        Mismatches:     27
Query Match:            28.72%        Indels:         15
DB:                     11            Gaps:           2

US-09-873-409-8 (1-514) x AK020318 (1-872)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          |||   |||||||||||||||||:::::||||||||    :::||||||||   |||
Db    230 ATGACTCTGGGAATATTAGCATCCATGATAAATGGAGCCACCGTTCCTTTAATGTCCCTG 289

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
          ||||||||||||:::||||||:::||||||:::||||||||||||||||||||
Db    290 GTTTTAGGAGAAATAAGTGATCATTTAATTAATGGATGCCTAGTACAAACTAACAGAACT 349

Qy     41 SerPhe------------------------------------------PheArgLeuThr 46
             :::                                                 ||||||
Db    350 AAATATCAGAACTGTTCTCAGACTCAAGAAAAGCTGAATGAAGATATCATTGTGTTGACT 409

Qy     47 LeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeu 66
          |||||||||:::||||||||   ||||||||||||||||||||:::|||||||||
Db    410 CTATATTATATTGGAATAGGAGCAGCTGCCCTCATTTTTGGCTATGTACAGATTTCCTTC 469

Qy     67 TrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerVal 86
          |||:::||||||||||||||||||||   |||||||||||||||||||||||||:::
Db    470 TGGGTCATAACTGCAGCCCGGCAAACCACAAGAATCCGAAAACAGTTTTTTCATTCAATT 529

Qy     87 LeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMet 106
          |||||||||||||||   |||||||||         ||||||   ||||||||||||||||||
Db    530 TTGGCACAAGACATCAGCTGGTTTGATGGCAGTGACATCTGTGAACTTAACACCCGCATG 589

Qy    107 Thr---AspIleAspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGln 125
          |||   ||||||:::|||:::   |||||||||||||||||||||   |||:::||||||
Db    590 ACTGGTGACATCAACAAACTCTGTGATGGTATTGGAGATAAGATCCCTCTGATGTTTCAG 649

Qy    126 AsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThr 145
          |||:::|||   |||||||||||||||    :::    |||:::|||   |||||||||:::
Db    650 AACATATCTGGGTTTTCTATTGGCCTGGTGATAAGTTTGATAAAAAGCTGGAAACTCTCC 709

Qy    146 LeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMet 165
          ||||||   ||||||||||||||||||||||||||||||:::|||   |||||||||||
```

```
Db      710 CTGGTGGTTCTGTCTACATCTCCTCTCATAATGGCTTCATCGGCACTGTGCTCTAGGATG 769

Qy      166 ValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGlu 185
            :::|||||||||||||||||||||||    |||||||||||||||||||||||||||||
Db      770 ATTATCTCATTGACCAGCAAGGAGCTGGATGCCTATTCCAAAGCTGGGGCTGTGGCTGAA 829

Qy      186 GluValLeuSerSerIleArgThrValIleAlaPheArgAla 199
            |||   |||||||||||:::||||||    ||||||   |||
Db      830 GAAGCCTTGTCATCAATCCAAACAGTCACAGCCTTTGGAGCC 871

RESULT 9
AK014319
LOCUS       AK014319                2676 bp    mRNA    linear   HTC 19-JAN-2002
DEFINITION  Mus musculus 14, 17 days embryo head cDNA, RIKEN full-length
            enriched library, clone:3222401P09:ATP-binding cassette, sub-family
            B (MDR/TAP), member 8, full insert sequence.
ACCESSION   AK014319
VERSION     AK014319.1  GI:12852089
KEYWORDS    HTC; CAP trapper.
SOURCE      Mus musculus (strain:C57BL/6J) 14, 17 days embryo head cDNA to
            mRNA, clone_lib:RIKEN full-length enriched mouse cDNA library
            clone:3222401P09.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1
  AUTHORS   Carninci,P. and Hayashizaki,Y.
  TITLE     High-efficiency full-length cDNA cloning
  JOURNAL   Meth. Enzymol. 303, 19-44 (1999)
  MEDLINE   99279253
   PUBMED   10349636
REFERENCE   2
  AUTHORS   Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
            Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
  TITLE     Normalization and subtraction of cap-trapper-selected cDNAs to
            prepare full-length cDNA libraries for rapid discovery of new genes
  JOURNAL   Genome Res. 10 (10), 1617-1630 (2000)
  MEDLINE   20499374
   PUBMED   11042159
REFERENCE   3
  AUTHORS   Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
            Konno,H., Akiyama,J., Nishi,K., Kitsunai,T., Tashiro,H., Itoh,M.,
            Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
            Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
            Fujiwake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
            Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
            Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
  TITLE     RIKEN integrated sequence analysis (RISA) system--384-format
            sequencing pipeline with 384 multicapillary sequencer
  JOURNAL   Genome Res. 10 (11), 1757-1771 (2000)
  MEDLINE   20530913
   PUBMED   11076861
REFERENCE   4
  AUTHORS   Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y.,
            Arakawa,T., Hara,A., Fukunishi,Y., Konno,H., Adachi,J., Fukuda,S.,
            Aizawa,K., Izawa,M., Nishi,K., Kiyosawa,H., Kondo,S., Yamanaka,I.,
            Saito,T., Okazaki,Y., Gojobori,T., Bono,H., Kasukawa,T., Saito,R.,
            Kadota,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T.,
            Fleischmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H.,
            Kuehl,P., Lewis,S., Matsuo,Y., Nikaido,I., Pesole,G.,
            Quackenbush,J., Schriml,L.M., Staubli,F., Suzuki,R., Tomita,M.,
            Wagner,L., Washio,T., Sakai,K., Okido,T., Furuno,M., Aono,H.,
            Baldarelli,R., Barsh,G., Blake,J., Boffelli,D., Bojunga,N.,
            Carninci,P., de Bonaldo,M.F., Brownstein,M.J., Bult,C.,
            Fletcher,C., Fujita,M., Gariboldi,M., Gustincich,S., Hill,D.,
            Hofmann,M., Hume,D.A., Kamiya,M., Lee,N.H., Lyons,P.,
            Marchionni,L., Mashima,J., Mazzarelli,J., Mombaerts,P., Nordone,P.,
            Ring,B., Ringwald,M., Rodriguez,I., Sakamoto,N., Sasaki,H.,
            Sato,K., Schonbach,C., Seya,T., Shibata,Y., Storch,K.F., Suzuki,H.,
            Toyo-oka,K., Wang,K.H., Weitz,C., Whittaker,C., Wilming,L.,
            Wynshaw-Boris,A., Yoshida,K., Hasegawa,Y., Kawaji,H., Kohtsuki,S.
            and Hayashizaki,Y.
  TITLE     Functional annotation of a full-length mouse cDNA collection
  JOURNAL   Nature 409 (6821), 685-690 (2001)
  MEDLINE   21085660
   PUBMED   11217851
REFERENCE   5  (bases 1 to 2676)
  AUTHORS   Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
            Arakawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bult,C.,
            Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,
            Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hiraoka,T., Hori,F.,
            Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,
            Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,
            Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
            Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,
            Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
            Schriml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
            Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
            Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I.,
            Yasunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M. and
            Hayashizaki,Y.
  TITLE     Direct Submission
  JOURNAL   Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of
            Physical and Chemical Research (RIKEN), Laboratory for Genome
            Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
            RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
            Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
            URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
            Fax:81-45-503-9216)
COMMENT     Please visit our web site (http://genome.gsc.riken.go.jp/) for
            further details.
            cDNA library was prepared and sequenced in Mouse Genome
            Encyclopedia Project of Genome Exploration Research Group in Riken
            Genomic Sciences Center and Genome Science Laboratory in RIKEN.
            Division of Experimental Animal Research in Riken contributed to
            prepare mouse tissues. First strand cDNA was primed with a primer
            [5' GAGAGAGAGAAGGATCCAAGAGCTCTTTTTTTTTTTTTTTTVN 3'], cDNA was
            prepared by using trehalose thermo-activated reverse transcriptase
            and subsequently enriched for full-length by cap-trapper. Second
            strand cDNA was prepared with the primer adapter of sequence [5'
            GAGAGAGAGATTCTCGAGTTAATTAAATTAATCCCCCCCCCCCC 3']. cDNA was cleaved
            with BamHI and XhoI. cDNA of size comprised between 3 and 7 kb was
            selected before cloning. Vector: a modified pBluescript KS(+) after
            bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'
            end: BamHI. Host: DH10B.
FEATURES             Location/Qualifiers
     source          1..2676
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="FANTOM_DB:3222401P09"
                     /db_xref="MGD:MGI:1907488"
                     /db_xref="taxon:10090"
                     /clone="3222401P09"
                     /tissue_type="head"
                     /clone_lib="RIKEN full-length enriched mouse cDNA library"
                     /dev_stage="14, 17 days embryo"
     gene            1..2676
                     /gene="Abcb8"
     CDS             72..2225
                     /gene="Abcb8"
                     /note="ATP-binding cassette, sub-family B (MDR/TAP),
                     member 8
                     data source:MGD, source key:MGI:1351667, evidence:ISS
                     putative"
                     /codon_start=1
                     /protein_id="BAB29270.1"
                     /db_xref="GI:12852090"
                     /db_xref="MGD:MGI:1351667"
                     /translation="MLVHLFRFGIRGGPVPGWSLQSLRFQTFSAARSSDDRLSSHLLR
                     TVAQLRVQLRAHLPRAPPASHWSPSAWCWVGGTLVVPAVLWQHPRLCLIALCEAKESP
                     PAQPTRAPELRFNWKLFWHFLHPHLLALGAAIVLALGAALVNVQIPLLLGQLVEIVAK
                     YTRDHMGSFVSESRKLSVQLLLLYGVQGLLTFGYLVLLSHIGERMAMDMRKALFSSLL
                     RQDIAFFDAKKTGQLVSRLTTDVQEFKSSFKLVISQGLRSCTQVIGSLVSLSMLSPRL
                     TLMLAVVTPALMGVGTLMGSGLRKLSRQCQEQIARATGVADEALGNVRTVRAFAMEKR
                     EEERYQAELESCCCKAEELGRGIALFQGLSNIAFNCMVLGTLFIGGSLVAGQQLKGGD
                     LMSFLVASQTVQRSMASLSVLFGQVVRGLSAGARVFEYMALSPVIPLTGGYCIPNKDI
```

```
                    RGSITFQNVTFSYPCRPGFNVLKDFTLKLPSGKIVALVGQSGGGKTTVASLLERFYDP
                    EAGSVTLDGHDLRTLNPSWLRGQVIGFISQEPVLFATTIMENIRFGKLDASDEEVYTA
                    AREANAHEFISSFPDGYSTVVGERGTTLSGGQKQRLAIARALIKQPTVLILDEATSAL
                    DAESERVVQEALDRASAGRTVLVIAHRLSTVRAAHSIIVMANGQVCEAGTHEELLKKG
                    GLYSELIRRQTLDASLTSTPPAEKPEDPKSCQSKA"
     polyA_signal   2653..2658
                    /gene="Abcb8"
                    /note="putative"
     polyA_site     2676
                    /gene="Abcb8"
                    /note="putative"
BASE COUNT      544 a    785 c    756 g    591 t
ORIGIN

Alignment Scores:
Pred. No.:              6.67e-73       Length:          2676
Score:                  710.00         Matches:         194
Percent Similarity:     54.48%         Conservative:    98
Best Local Similarity:  36.19%         Mismatches:      184
Query Match:            27.39%         Indels:          60
DB:                     11             Gaps:            19

US-09-873-409-8 (1-514) x AK014319 (1-2676)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
             ::::::|||   :::    |||:::||||||||              :::||||||
Db     471 ATTGTGCTGGCCTTAGGTGCGGCACTAGTGAAT------------GTGCAGATCCCCTTG 518

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIleSer------------------GlyCysLeu 34
             :::|||||||:::::   :::   :::                        |||
Db     519 CTCCTGGGCCAGCTGGTGGAGATTGTCGCCAAGTACACGAGGGACCACATGGGGAGTTTC 578

Qy      35 ValGlnThrAsnThrTyrSerPhePheArgLeuThrLeuTyrTyrValGlyIleGlyVal 54
             |||      :::               :::|||  |||  |||         ||||||
Db     579 GTGTCTGAGTCCCGTAAGCTCAGCGTCCAGCTGCTCCTACTGTAC---------GGTGTT 629

Qy      55 AlaAlaLeuIle---PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArg 73
                   |||:::   |||||||||:::   :::                    |||
Db     630 CAGGGACTGCTGACCTTTGGATACCTAGTGCTGCTGTCCCACATT------GGTGAGCGC 683

Qy      74 GlnThrLysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrp 93
                       :::||||||     |||   |||:::|||   |||||||||   :::
Db     684 ATGGCCATGGACATGCGGAAAGCCCTTTTCAGCTCCCTGCTCCGGCAAGACATTGCTTTC 743

Qy      94 PheAspSerCysAspIleGlyGluLeuAsnThrArgMet---ThrAspIleAspLys--- 111
             ||||||:::          |||:::|||   :::|||:::   ||||||:::   :::
Db     744 TTTGATGCCAAAAAGACAGGGCAGCTAGTGAGTCGCTTGACTACTGATGTGCAAGAATTC 803

Qy     112 ---------------------IleSerAspGlyIleGlyAspLysIleAlaLeuLeuPhe 124
                                 ||||||   |||:::
Db     804 AAGTCATCCTTCAAGCTTGTCATCTCCCAGGGACTG------------------------ 839

Qy     125 GlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeu---ValLysGlyTrpLys 143
                       |||      ||||||   |||   |||      :::      :::
Db     840 ---CGCAGCTGCACCCAGGTGATTGGTAGCCTGGTGTCCCTGTCTATGCTGTCCCCTCGC 896

Qy     144 LeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSer 163
             |||||||||:::                :::|||   :::|||
Db     897 CTTACCCTGATGCTGGCTGTCGTCACACCCGCCCTCATGGAGTGGGCACCCTGATGGGC 956

Qy     164 ArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaVal 183
                   :::      |||:::   :::                  ::::::|||      |||
Db     957 TCAGGCCTCCGAAAGCTGTCTCGCCAGTGTCAGGAGCAGATTGCCAGGGCAACAGGTGTA 1016

Qy     184 AlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGlu 203
             |||:::|||   |||   ::::::|||||||||   ||||||         :::::::::|||
Db    1017 GCAGATGAGGCCCTTGGCAATGTTCGGACTGTGCGGGCCTTCGCGATGGAGAAGAGGGAG 1076

Qy     204 LeuGlnArgTyrThrGlnAsnLeuLysAsp------------AlaLysAspPheGlyIle 219
                :::||||||         |||:::                     |||::::::   |||
Db    1077 GAGGAACGCTATCAAGCAGAACTGGAGTCATGCTGCTGTAAAGCAGAAGAACTGGGC--- 1133

Qy     220 LysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhe------PheMetAsn 237
                |||   ||||||                |||      :::   |||               :::
Db    1134 ---AGGGGCATCGCCTTGTTCCAAGGGCTCTCCAACATCGCTTTCAACTGTATGGTCTTG 1190

Qy     238 GlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGlu---Pro 256
             ||||||                   :::   |||   ||||||:::         :::
Db    1191 GGCACC------------CTGTTCATTGGGGGCTCCCTTGTGGCTGGACAGCAGCTGAAA 1238

Qy     257 GlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCys 276
             |||      :::   :::   |||                  :::|||         |||
Db    1239 GGGGAGACCTCATGTCCTTCCTGGTGGCTTCCCAGACAGTACAGAGGTCTATGGCCAGC 1298

Qy     277 IleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGly------AlaAla 294
             :::                                    :::   ||||||      |||
Db    1299 CTCTCTGTCCTGTTTGGTCAG---------------GTGGTACGTGGGCTGAGTGCGGGA 1343

Qy     295 PheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGly 314
                   :::|||:::   :::               |||   |||               |||   |||
Db    1344 GCCCGAGTCTTCGAATACATGGCCCTGAGCCCTGTCATCCCA------TTGACCGGGGGC 1397

Qy     315 Tyr------LysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsn 332
             |||               :::   |||   |||::::::   |||:::|||||:::|||:::
Db    1398 TACTGCATCCCCAACAAGGACATTCGTGGTTCCATCACCTTCCAAAATGTCACCTTCAGC 1457

Qy     333 TyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGly 352
             ||||||   ||||||            :::||||||         |||::::::   ||||||
Db    1458 TACCCCTGCAGACCTGGCTTCAATGTGCTCAAGGACTTCACCCTGAAGCTGCCCTCTGGC 1517

Qy     353 GluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeu 372
             :::   |||||||||||||||   :::|||   ||||||:::||||||         ||||||
Db    1518 AAGATTGTGGCTCTTGTGGGCCAGTCTGGGGGAGGAAAGACCACAGTTGCCTCCTTGCTG 1577

Qy     373 GlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAla 392
             :::|||   |||||||||:::   |||   :::   :::|||   :::|||:::|||
Db    1578 GAACGCTTCTATGACCCTGAAGCTGGCTCGGTGACGTTGGATGGCATGACCTGCGAACT 1637

Qy     393 LeuAsnValArgHisTyrArgAspHis---IleGlyValValSerGlnGluProValLeu 411
             ||||||         |||         ||||||   :::||||||||||||||||||
Db    1638 CTCAACCCCTCCTGGCTCCGGGGCCAGGTCATAGGTTTCATCAGCCAGGAGCCAGTCCTG 1697

Qy     412 PheGlyThrThrIleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGlu 431
             |||   |||||||||         ||||||:::::::::|||   |||   :::|||||||||
Db    1698 TTTGCAACAACCATCATGGAGAATATCCGATTTGGGAAGCTGGATGCTTCCGATGAAGAG 1757

Qy     432 MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLys 451
             :::   ||||||||||||||||||||||||::::::||||||         ||||||:::
Db    1758 GTGTACACAGCTGCACGAGAAGCCAATGCCCACGAGTTCATCAGCAGCTTCCCCGATGGC 1817

Qy     452 PheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIle 471
             ::::::|||:::|||||||||:::|||         :::||||||||||||||||||:::
Db    1818 TACAGCACTGTGGTTGGTGAGCGGGGCACAACCTTGTCTGGTGGCCAGAAGCAGCGCCTA 1877

Qy     472 AlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSer 491
             ||||||||||||||||||::::::   |||   :::||||||||||||||||||||||||
Db    1878 GCCATCGCACGTGCCCTCATCAAGCAGCCCACAGTGCTGATCCTGGACGAGGCCACCAGT 1937

Qy     492 AlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 507
             |||||||||:::||||||:::         ||||||   ||||||::::::
Db    1938 GCGCTAGATGCAGAATCCGAGAGGGTGGTACAGGAGGCCCTGGACCGG 1985

RESULT 10
BG298756
LOCUS       BG298756               803 bp    mRNA    linear   EST 21-FEB-2001
DEFINITION  602396681F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:4511157 5',
            mRNA sequence.
ACCESSION   BG298756
VERSION     BG298756.1  GI:13063728
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 803)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
```

```
  TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL     Unpublished (1999)
COMMENT       Contact: Robert Strausberg, Ph.D.
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: The Cepko Laboratory
               cDNA Library Preparation: Life Technologies, Inc.
               cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
               DNA Sequencing by: Incyte Genomics, Inc.
               Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LLNL at:
              http://image.llnl.gov
              Plate: LLAM10394  row: b  column: 22
              High quality sequence stop: 698.
FEATURES               Location/Qualifiers
     source            1..803
                       /organism="Mus musculus"
                       /db_xref="taxon:10090"
                       /clone="IMAGE:4511157"
                       /clone_lib="NIH_MGC_94"
                       /tissue_type="retina"
                       /lab_host="DH10B (phage-resistant)"
                       /note="Organ: eye; Vector: pCMV-SPORT6; Site_1: NotI;
                       Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
                       Average insert size 3.3 kb. Library enriched for
                       full-length clones and constructed by Life Technologies.
                       Note: this is a NIH_MGC Library."
BASE COUNT      240 a    162 c    205 g    196 t
ORIGIN

Alignment Scores:
Pred. No.:              3.63e-69       Length:          803
Score:                  671.50         Matches:         149
Percent Similarity:     76.15%         Conservative:    49
Best Local Similarity:  57.31%         Mismatches:      55
Query Match:            25.91%         Indels:          10
DB:                     12             Gaps:            2

US-09-873-409-8 (1-514) x BG298756 (1-803)

Qy     168 SerLeuThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluVal 187
           |||   |||   ||||||||   |||||:::||||||||||||||||||||||||
Db      15 TCATTTACTGATAAGGAACTCCATGCTTATGCAAAAGCTGGAGCAGTTGCTGAAGAAGTC 74

Qy     188 LeuSerSerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgTyr 207
           |||::::::|||||||||||||||||             |||:::|||||||||:::||||||
Db      75 TTAGCAGCCATCAGAACTGTGATTGCGTTGGA-GGACAAAAGAAGGAACTTGAAAGGTAC 133

Qy     208 ThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysVal 227
                 |||||:::::||||||      ||||||||:::   |||    :::   :::
Db     134 AATAACAACTTGGAAGAAGCTAAAAGGCTGGGGATAAAGAAAGCTATCACGGCCAACATC 193

Qy     228 SerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGly 247
           |||:::||||||   :::      :::      :::|||   ||||||||||||||||||
Db     194 TCCATGGGTGCAGCTTTTCTCCTTATCTATGCATCATATGCTCTGGCATTCTGGTATGGG 253

Qy     248 ThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhe 267
           ||||||||:::::::::   |||         |||:::   |||   ||||||   ||||||
Db     254 ACTTCCTTGGTCATCTCCAAAGAA------TACTCTAT-GGACAAGTGCTCACTGTCTTC 306

Qy     268 PheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPhe 287
           |||||||||:::      ::::::   :::|||   |||   |||:::   |||   |||
Db     307 TTTTCCGTGTTAATTGGAGCATTCAGTGTTGGACAGGCATCTCCAAATATTGAAGCCTTC 366

Qy     288 AlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIle 307
           |||   ||||||||||||||||:::   :::|||::::::||||||   ||||||||||||
Db     367 GCCAATGCACGAGGAGCAGCTTATGAAGTCTTCAAAATAATTGATAATAAGCCCAGTATA 426

Qy     308 AspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLys 327
           |||:::||||||   :::|||:::||||||::::::|||:::|||   :::|||||||||
Db     427 GACAGCTTCTCAAAGAGTGGGCACAAACCAGACAACATACAAGGAAATCTGGAATTTAAG 486

Qy     328 AsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeu 347
           |||:::   |||:::||||||||||||      ::::::||||||||||||||||||||
Db     487 AATATTCACTTCAGTTACCCATCTCGAAAAGAAGTTCAGATCTTGAAGGGCCTCAATCTG 546

Qy     348 ArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThr 367
           ::::::||||||||:::||||||||||||||||||   :::|||   |||||||||||
Db     547 AAGGTGAAGAGCGGACAGACGGTGGCCCTGGTTGGCAACAGTGGCTGTGGAAAAAGCACA 606

Qy     368 ValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGlu 387
           |||||||||:::||||||||||||||||||   ||||||   :::   :::|||
Db     607 ACTGTCCAGCTGATGCAAAGGCTCTACGACCCCTA-GATGGCATGGTCAGTATCGACGGA 665

Qy     388 AsnAspIleArgAlaLeu-AsnValArgHis-TyrArg-AspHisIleGlyValValSer 406
           |||||||||   :::   |||||||||:::   :::|||   ||||||   ||||||||||||
Db     666 CAGGACATCAGAACCATCCAATGTGAGGTATCTGGAGGGGATCAT---GGGGTGGTGAGT 722

Qy     407 Gln-GluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 423
           |||   ||||||   ||||||               |||:::||||||
Db     723 CAAGGAACCTGGCTTGTTTGCCAACCCAGATCGCGAGAAAATTCGTTATGGG 774

RESULT 11
AL520322
LOCUS       AL520322                 943 bp    mRNA    linear   EST 13-FEB-2001
DEFINITION  AL520322 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB006YC15 5
            prime, mRNA sequence.
ACCESSION   AL520322
VERSION     AL520322.1  GI:12783815
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 943)
  AUTHORS   Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
  TITLE     Full-length cDNA libraries and normalization
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES             Location/Qualifiers
     source          1..943
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CS0DB006YC15"
                     /clone_lib="LTI_NFL004_NBC2"
                     /sex="male"
                     /tissue_type="neuroblastoma cells"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
                     was primed with a NotI-oligo(dT) primer. Five prime end
                     enriched, double-stranded cDNA was digested with Not I and
                     cloned into the Not I and Eco RV sites of the pCMVSPORT 6
                     vector.  Library was normalized. Library was constructed
                     by Life Technologies.  Contact : Feng Liang Life
                     Technologies, a division of Invitrogen 9800 Medical Center
                     Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
                     8371 Email : fliang@lifetech.com URL :
                     http://fulllength.invitrogen.com"
BASE COUNT      253 a    225 c    253 g    211 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:              1.47e-67       Length:          943
Score:                  659.00         Matches:         131
Percent Similarity:     72.36%         Conservative:    47
Best Local Similarity:  53.25%         Mismatches:      66
Query Match:            25.42%         Indels:          2
DB:                     9              Gaps:            1

US-09-873-409-8 (1-514) x AL520322 (1-943)

Qy     264 LeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHis 283
           |||   :::|||   :::|||:::         :::      :::|||
Db       2 CTGTTATTATTTTCAGCTGTTGTCTTTGGTGCCATGGCCGTGGGGCAAGTCAGTTCATTT 61
```

```
Qy     284 PheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLys 303
               :::|||   |||:::   :::|||  ||||||     :::|||:::|||
Db      62 GCTCCTGACTATGCCAAAGCCAAAATATCAGCAGCCCACATCATCATGATCATTGAAAAA 121

Qy     304 LysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThr 323
              |||   ||||||::::::||||||   |||      |||    ::::::||||||
Db     122 ACCCCTTTGATTGACAGCTACAGCACGGAAGGCCTAATGCCGAACACATTGGAAGGAAAT 181

Qy     324 ValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLys 343
           |||   |||      |||   ||||||||||||:::||||||   |||   :::|||:::
Db     182 GTCACATTTGGTGAAGTTGTATTCAACTATCCCACCCGACCGGACATCCCAGTGCTTCAG 241

Qy     344 GlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySer 363
           ||||||:::|||   :::|||   |||:::||||||||||||      :::|||
Db     242 GGACTGAGCCTGGAGGTGAAGAAGGGCCAGACGCTGGCTCTGGTGGGCAGCAGTGGCTGT 301

Qy     364 GlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIle 383
           ||||||||||||||||||||||||||||:::|||   |||||||||      |||   :::
Db     302 GGGAAGAGCACAGTGGTCCAGCTCCTGGAGCGGTTCTACGACCCCTTGCAGGGAAAGTG 361

Qy     384 MetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHisIleGly 403
           ::::::|||      :::|||:::   |||||||||:::      |||   |||:::|||
Db     362 CTGCTTGATGGCAAAGAAATAAAGCGACTGAATGTTCAGTGGCTCCGAGCACACCTGGGC 421

Qy     404 ValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly 423
           :::|||||||||||||||:::||||||      :::|||:::   ||||||   ||||||
Db     422 ATCGTGTCCCAGGAGCCCATCCTGTTTGACTGCAGCATTGCTGAGAACATTGCCTATGGA 481

Qy     424 ArgAspAsp------ValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla 441
              :::         |||:::   ||||||:::   ||||||||:::|||||||||
Db     482 GACAACAGCCGGGTGGTGTCACAGGAAGAGATCGTGAGGGCAGCAAAGGAGGCCAACATA 541

Qy     442 TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAla 461
           :::   ||||||            |||||||||::::::|||   ||||||:::||||||
Db     542 CATGCCTTCATCGAGTCACTGCCTAATAAATATAGCACTAAAGTAGGAGACAAAGGAACT 601

Qy     462 GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro 481
           |||:::||||||||||||||||||||||||||||||||||||||||||||||   |||
Db     602 CAGCTCTCTGGTGGCCAGAAACAACGCATTGCCATAGCTCGTGCCCTTGTTAGACAGCCT 661

Qy     482 LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 501
              ||||||:::||||||||||||||||||||||||||||:::||||||:::      |||
Db     662 CATATTTTGCTTTTGGATGAAGCCACGTCAGCTCTGGATACAGAAAGTGAAAAGGTTGTC 721

Qy     502 GlnAlaAlaLeuGluLys 507
           |||   ||||||:::|||
Db     722 CAAGAAGCCCTGGACAAA 739
```

RESULT 12
AA243820

```
LOCUS       AA243820                 405 bp    mRNA    linear   EST 06-AUG-1997
DEFINITION  zr67g06.r1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:668506 5'
            similar to SW:MDR1_MOUSE P06795 MULTIDRUG RESISTANCE PROTEIN 1 ;,
            mRNA sequence.
ACCESSION   AA243820
VERSION     AA243820.1  GI:1874631
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 405)
  AUTHORS   Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
            Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J., Moore,B.
            , Schellenberg,K., Steptoe,M., Tan,F., Theising,B., White,Y., Wylie
            ,T., Waterston,R. and Wilson,R.
  TITLE     WashU-Merck EST Project 1997
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1572   Std Error: 0.00
            Seq primer: -28m13 rev2 ET from Amersham.
FEATURES             Location/Qualifiers
     source          1..405
                     /organism="Homo sapiens"
                     /db_xref="GDB:5562473"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:668506"
                     /clone_lib="Soares_NhHMPu_S1"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."
BASE COUNT      115 a     81 c     85 g    124 t
ORIGIN

Alignment Scores:
Pred. No.:               1.44e-67       Length:          405
Score:                   654.00         Matches:         128
Percent Similarity:      100.00%        Conservative:    1
Best Local Similarity:   99.22%         Mismatches:      0
Query Match:             25.23%         Indels:          0
DB:                      9              Gaps:            0

US-09-873-409-8 (1-514) x AA243820 (1-405)

Qy     264 LeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAlaValProHis 283
           |||:::||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      18 CTCTCTGTTTTCTTTAGTGTAATCCATAGCAGTTATTGCATTGGAGCAGCAGTCCCTCAC 77

Qy     284 PheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLys 303
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      78 TTTGAAACCTTCGCAATAGCCCGAGGAGCTGCCTTTCATATTTTCCAGGTTATTGATAAG 137

Qy     304 LysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIleGluGlyThr 323
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     138 AAACCCAGTATAGATAACTTTTCCACAGCTGGATATAAACCTGAATCCATAGAAGGAACT 197

Qy     324 ValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIleLeuLys 343
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     198 GTGGAATTTAAAAATGTTTCTTTCAATTATCCATCAAGACCATCTATCAAGATTCTGAAA 257

Qy     344 GlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySer 363
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     258 GGTCTGAATCTCAGAATTAAGTCTGGAGAGACAGTCGCCTTGGTCGGTCTCAATGGCAGT 317

Qy     364 GlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIle 383
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     318 GGGAAGAGTACGGTAGTCCAGCTTCTGCAGAGGTTATATGATCCGGATGATGGCTTTATC 377

Qy     384 MetValAspGluAsnAspIleArgAla 392
           |||||||||||||||||||||||||||
Db     378 ATGGTGGATGAGAATGACATCAGAGCT 404
```

RESULT 13
BJ075336

```
LOCUS       BJ075336                 550 bp    mRNA    linear   EST 11-DEC-2001
DEFINITION  BJ075336 NIBB Mochii normalized Xenopus tailbud library Xenopus
            laevis cDNA clone XL057d10 5', mRNA sequence.
```

```
ACCESSION   BJ075336
VERSION     BJ075336.1  GI:17520252
KEYWORDS    EST.
SOURCE      African clawed frog.
  ORGANISM  Xenopus laevis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus.
REFERENCE   1  (bases 1 to 550)
  AUTHORS   Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-i,T. and Kohara
            ,Y.
  TITLE     Expressed genes in X. laevis embryo
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp.
FEATURES             Location/Qualifiers
     source          1. .550
                     /organism="Xenopus laevis"
                     /db_xref="taxon:8355"
                     /clone="XL057d10"
                     /clone_lib="NIBB Mochii normalized Xenopus tailbud
                     library"
                     /tissue_type="whole embryo"
                     /dev_stage="stage 25"
                     /note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs
                     were oligo-dT primed and directionally cloned. Staging
                     according to  Nieuwkoop and Faber. Library is subtracted
                     and was constructed by N. Garrett  and A.M. Zorn,
                     (Wellcome/CRC Institute). "
BASE COUNT      162 a    109 c    144 g    135 t
ORIGIN

Alignment Scores:
Pred. No.:              7.17e-67      Length:         550
Score:                  650.00        Matches:        121
Percent Similarity:     85.79%        Conservative:   36
Best Local Similarity: .66.12%        Mismatches:     26
Query Match:            25.08%        Indels:         0
DB:                     13            Gaps:           0

US-09-873-409-8 (1-514) x BJ075336 (1-550)

Qy     322 GlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLysIle 341
           |||   :::|||||||||    :::    |||   ||||||   ||||||    ::::::|||
Db       2 GGCGATATAGAATTCAAAGGCATTCATTTTGCCTACCCATGCCGTCCTGGTGTGCAAATT 61

Qy     342 LeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeuAsn 361
           |||   ||||||||||||:::::::::::|||:::|||:::||||||||||||::::::
Db      62 CTCTCTGGACTGAACCTGAAAGTGGAAGCTGGTAAAACCATCGCTTTGGTTGGTATGAGT 121

Qy     362 GlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAspGly 381
           |||   ||||||||||||    :::||||||||||||||||   |||||||||       |||
Db     122 GGTTGTGGTAAAAGCACTACCATACAGTTGCTGCAGAGGTTCTATGATCCATCACATGGA 181

Qy     382 PheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAspHis 401
              :::   :::|||    :::|||||||||:::|||||||||:::         |||::::::
Db     182 GAGGTTACTTTGGATGGACACGATATCCGTTCACTAAATGTAAAATGGTTGAGGGAAAAT 241

Qy     402 IleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIleLys 421
           ||||||||||||||||||||||||||||||||||||||||||:::    ||||||:::
Db     242 ATTGGCGTAGTGAGTCAGGAGCCAGTGCTCTTTGGCACAACCATTGCAGAAAATATCCGC 301

Qy     422 TyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAla 441
           ||||||    :::    |||||||||   |||:::|||:::||||||:::|||||||||||
Db     302 TACGGAGGGGAATCAGTAACAGACAACGAAATTGAACAAGCGGCAAAAGAAGCCAATGCC 361

Qy     442 TyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAla 461
           :::|||||||||          |||:::||||||   |||:::|||||||||:::||||||
Db     362 TTTGATTTTATATCCAGGCTTCCTGATAAATTTAAGACGATGGTGGGAGAAAGAGGTGCG 421

Qy     462 GlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnPro 481
           |||:::||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     422 CAGTTGAGTGGGGGTCAGAAGCAGCGGATAGCCATTGCACGCGCCCTGGTACGGAACCCT 481

Qy     482 LysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaVal 501
           |||||||||:::|||||||||||||||||||||||||||||::::::|||::::::    |||
Db     482 AAGATTCTACTTCTTGATGAAGCCACATCAGCTCTGGATACACAAAGTGAAGCTATTGTA 541

Qy     502 GlnAlaAla 504
           |||||||||
Db     542 CAAGCTGCC 550

RESULT 14
BG587938
LOCUS       BG587938                 780 bp    mRNA    linear   EST 11-APR-2001
DEFINITION  EST489713 MHAM Medicago truncatula/Glomus versiforme mixed EST
            library cDNA clone pMHAM-51L19 5' end, mRNA sequence.
ACCESSION   BG587938
VERSION     BG587938.1  GI:13603002
KEYWORDS    EST.
SOURCE      Medicago truncatula/Glomus versiforme mixed EST library.
  ORGANISM  Medicago truncatula/Glomus versiforme mixed EST library
            Eukaryota; mixed EST libraries.
REFERENCE   1  (bases 1 to 780)
  AUTHORS   Harrison,M.J., Liu,J., Town,C.D., Van Aken,S., Utterback,T., Cho,J.
            and Fraser,C.M.
  TITLE     ESTs from roots of Medicago truncatula after colonization with
            Glomus versiforme, 2001
  JOURNAL   Unpublished (2001)
COMMENT     Contact: Harrison M.J.
            Plant Biology Division
            The Samuel Roberts Noble Foundation
            2510 Sam Noble Parkway, Ardmore, OK 73401
            Tel: 580-223-5810
            Fax: 580-221-7380
            Email: mjharrison@noble.org
            Noble EST name: N387011e TIGR sequence name: MTDDM70TK More
            information is available at: http://www.medicago.org
            Seq primer: SKmod (CTA gAA CTA gtg gAT CC).
FEATURES             Location/Qualifiers
     source          1. .780
                     /organism="Medicago truncatula/Glomus versiforme mixed EST
                     library"
                     /cultivar="Medicago truncatula genotype A17"
                     /db_xref="taxon:119092"
                     /clone="pMHAM-51L19"
                     /clone_lib="MHAM"
                     /tissue_type="roots colonized with Glomus versiforme"
                     /dev_stage="Roots harvested at 10, 17, 22, 31 and 38 days
                     post-inoculation with Glomus versiforme. The library was
                     made from a mixture of RNA from each of these stages."
                     /lab_host="E. coli strain XLOLR"
                     /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
                     XhoI; cDNA was prepared from polyA+ enriched RNA from
                     roots harvested at 10, 17, 22, 31 and 38 days
                     post-inoculation with Glomus versiforme. The cDNA was
                     directionally ligated into the Unizap XR vector from
                     Stratagene and packaged using Gigapack III Gold packaging
                     extracts. Plasmids containing cDNA inserts were excised
                     from the recombinant lambda-Zap phage using Ex-assist
                     helper phage and propagated in XLOLR cells."
BASE COUNT      234 a    156 c    184 g    206 t
ORIGIN

Alignment Scores:
Pred. No.:              4.53e-65      Length:         780
Score:                  637.00        Matches:        126
Percent Similarity:     71.43%        Conservative:   39
Best Local Similarity:  54.55%        Mismatches:     66
Query Match:            24.58%        Indels:         0
DB:                     12            Gaps:           0
```

```
US-09-873-409-8 (1-514) x BG587938 (1-780)

Qy    277 IleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHis 296
           :::|||   :::   |||              |||    ||||||   ||||||
Db      3 TTGGGGCAATCTGCACCGAGCATGGCTGCATTTACAAAGGCTAGAGTTGCGGCTGCTAAG 62

Qy    297 IlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLys 316
          ||||||:::::||||||     :::|||   ||||||         |||   :::|||   :::
Db     63 ATTTTCCGGATAATTGATCACCAGCCTGGTATAGATAGAAACAGTGAATCTGGATTGGAA 122

Qy    317 ProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArg 336
                |||::::::   |||   ||||||   ||||||||   |||:::||||||||||||
Db    123 TTAGAGACAGTTACTGGACTTGTTGAACTGAAAAATGTGGACTTCTCTTATCCATCAAGA 182

Qy    337 ProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAla 356
          |||   :::   ||||||         :::|||   :::   :::|||:::|||:::|||
Db    183 CCTGAAGTTCTGATCCTCAATGATTTCTCCTTGAGTGTTCCTGCCGGAAAGACCATAGCT 242

Qy    357 LeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyr 376
          |||||||||   :::||||||||||||||||||||||||   |||:::::::|||   |||
Db    243 TTAGTTGGTAGCAGCGGCTCTGGCAAGAGCACTGTTGTTTCTCTTATTGAGAGATTCTAT 302

Qy    377 AspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArg 396
          ||||||      |||   :::|||:::|||      :::||||||:::   |||   ::::::
Db    303 GATCCAACTTCAGGACAAGTAATGTTGGATGGGCATGACATTAAAACTTTGAAGCTTAAA 362

Qy    397 HisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIle 416
                |||      ||||||:::|||||||||||||||   ||||||   |||||||||
Db    363 TGGTTGAGGCAACAAATAGGACTAGTGAGCCAAGAACCTGCTTTGTTTGCCACCACGATT 422

Qy    417 SerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAla 436
                ||||||      ||||||   |||               |||:::|||   ||||||
Db    423 CGAGAAAATATACTCTTGGGAAGGCCTGATGCAAACCAGGTTGAGATTGAAGAAGCTGCT 482

Qy    437 ArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuVal 456
          |||   |||||||||:::   ||||||::::::   |||      |||   |||   |||
Db    483 AGGGTTGCTAATGCTCATTCATTCATCATCAAACTTCCTGAAGGCTTTGAAACTCAGGTA 542

Qy    457 GlyGluLysGlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAla 476
          ||||||:::|||   |||:::|||||||||||||||||||||||||||||||||||||||
Db    543 GGAGAAAGAGGACTACAACTTTCTGGAGGACAAAAACAAAGAATAGCAATAGCAAGGGCA 602

Qy    477 LeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGlu 496
          :::::::::||||||   ||||||:::|||||||||||||||||||||||||||||||||
Db    603 ATGCTAAAAAATCCAGCAATTCTTCTCCTCGATGAGGCAACAAGTGCATTGGACTCTGAA 662

Qy    497 SerLysSerAlaValGlnAlaAlaLeuGluLys 507
          |||:::      ||||||   ||||||::::::
Db    663 TCAGAAAAGCTGGTGCAAGAAGCACTTGACCGG 695

RESULT 15
BF692596
LOCUS       BF692596                  559 bp    mRNA    linear   EST 22-DEC-2000
DEFINITION  602248949F1 NIH_MGC_62 Homo sapiens cDNA clone IMAGE:4334100 5',
            mRNA sequence.
ACCESSION   BF692596
VERSION     BF692596.1  GI:11978004
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 559)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: CLONETECH Laboratories, Inc.
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLCM1203  row: a  column: 13
            High quality sequence stop: 555.
FEATURES             Location/Qualifiers
     source          1. .559
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4334100"
                     /clone_lib="NIH_MGC_62"
                     /tissue_type="melanotic melanoma, high MDR"
                     /lab_host="DH10B (T1 phage-resistant)"
                     /note="Organ: skin; Vector: pDNR-LIB (Clontech); Site_1:
                     SfiI (ggccgcctcggcc); Site_2: SfiI (ggccattatggcc);
                     Double-stranded cDNA was prepared from cell line RNA.  5'
                     and 3' adaptors were used in cloning as follows: 5'
                     adaptor sequence:  5'-CACGGCCATTATGGCC-3' and 3' adaptor
                     sequence:  5'-ATTCTAGAGGCCGAGGCGGCCGACATG-dT(30)BN-3'
                     (where B = A, C, or G and N = A, C,  G, or T).   Average
                     insert size 1.75 kb (range 0.9-4.0 kb).  15/15 colonies
                     contained  inserts by PCR. This library was enriched for
                     full-length clones and was constructed by Clontech
                     Laboratories (Palo Alto, CA)."
BASE COUNT      161 a    118 c    141 g    139 t
ORIGIN

Alignment Scores:
Pred. No.:               7.08e-64        Length:          559
Score:                   625.00          Matches:         134
Percent Similarity:      96.43%          Conservative:    1
Best Local Similarity:   95.71%          Mismatches:      4
Query Match:             24.11%          Indels:          3
DB:                      12              Gaps:            0

US-09-873-409-8 (1-514) x BF692596 (1-559)

Qy    376 TyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVal 395
          ||||||||||||||||||||||||   |||||||||||||||||||||||||||||||||
Db      1 TATGATCCGGATGATGGCTTTATCACGGTGGATGAGAATGACATCAGAGCTTTAAATGTG 60

Qy    396 ArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThr 415
          ||||||||||||||||||   ||||||||||||||||||||   ||||||||||||||
Db     61 CGGCATTATCGAGACCATAT-GGAGTGGTTAGTCAAGAGCCTGT-TTGTTCGGGACCACC 118

Qy    416 IleSerAsnAsnIleLysTyrGlyArgAspAspValThrAspGluGluMetGluArgAla 435
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    119 ATCAGTAACAATATCAAGTATGGACGAGATGATGTGACTGATGAAGAGATGGAGAGAGCA 178

Qy    436 AlaArgGluAlaAsnAlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeu 455
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    179 GCAAGGGAAGCAAATGCATATGATTTTATCATGGAGTTTCCTAATAAATTTAATACATTG 238

Qy    456 ValGlyGluLys-GlyAlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaAr 475
          |||||||||     ||||||||||||||||||||||||||||||||||||||||||||||
Db    239 GTAGGGGAAAACAGGAGCTCAAATGAGTGGAGGGCAGAAACAGAGGATCGCAATTGCTCG 298

Qy    475 gAlaLeuValArgAsnProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSe 495
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    299 TGCCTTAGTTCGAAACCCCAAGATTCTGATTTTAGATGAGGCTACGTCTGCCCTGGATTC 358

Qy    495 rGluSerLysSerAlaValGlnAlaAlaLeuGluLysAspThrProArgTyrSerPhe 514
          |||||||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db    359 AGAAAGCGAGTCAGCTGTTCAAGCTGCACTGGAGAAGGATACCCCCAGGTATTCATTT 416

Search completed: March 31, 2003, 13:51:49
Job time : 3547.51 secs
```

```
                  GenCore version 5.1.4_p5_4578
          Copyright (c) 1993 - 2003  Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on:        March 31, 2003, 02:24:20 ; Search time 75.2727 Seconds
                                          (without alignments)
                                          2094.145 Million cell updates/sec

Title:          US-09-873-409-8
Perfect score:  2592
Sequence:       1 MILGILASLVNGACLPLMPL..........SESKSAVQAALEKDTPRYSF 514

Scoring table:  BLOSUM62
                Xgapop 10.0 , Xgapext  0.5
                Ygapop 10.0 , Ygapext  0.5
                Fgapop  6.0 , Fgapext  7.0
                Delop   6.0 , Delext   7.0

Searched:       441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters:      882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09873409/runat_27032003_115422_19259/app_query.fasta_1.7544
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09873409_@CGN_1_1_578_@runat_27032003_115422_19259 -NCPU=6 -ICPU=3
-NO_XLPXY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :      Issued_Patents_NA:*
                1:  /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
                2:  /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
                3:  /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
                4:  /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
                5:  /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
                6:  /cgn2_6/ptodata/2/ina/backfiles1.seq:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.
```

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	1576	60.8	4233	3	US-09-120-513-1	Sequence 1, Appli
	2	1576	60.8	4233	4	US-09-450-105-1	Sequence 1, Appli
	3	1576	60.8	4669	2	US-08-752-447-1	Sequence 1, Appli
	4	1576	60.8	4669	4	US-09-316-167-1	Sequence 1, Appli
	5	1570	60.6	4646	1	US-08-181-471-2	Sequence 2, Appli
	6	1570	60.6	6505	2	US-08-793-610-5	Sequence 5, Appli
	7	1570	60.6	9318	2	US-08-793-610-6	Sequence 6, Appli
	8	1568.5	60.5	4264	2	US-08-784-649A-1	Sequence 1, Appli
	9	1568.5	60.5	4264	2	US-08-784-649A-5	Sequence 5, Appli
	10	1568	60.5	4669	6	5206352-3	Patent No. 5206352
	11	1549	59.8	4669	2	US-08-583-276-18	Sequence 18, Appl
	12	975.5	37.6	4224	1	US-08-612-521-1	Sequence 1, Appli
	13	959	37.0	4047	2	US-08-612-734B-1	Sequence 1, Appli
	14	939.5	36.2	4800	2	US-08-612-734B-3	Sequence 3, Appli
	15	937.5	36.2	4002	2	US-08-996-545-1	Sequence 1, Appli
	16	937.5	36.2	4002	2	US-08-996-545-3	Sequence 3, Appli
	17	937.5	36.2	4002	4	US-09-328-320-1	Sequence 1, Appli
	18	937.5	36.2	4002	4	US-09-328-320-3	Sequence 3, Appli
	19	920	35.5	3924	1	US-08-395-246C-1	Sequence 1, Appli
	20	902.5	34.8	2726	1	US-08-461-823-1	Sequence 1, Appli
	21	859	33.1	6143	1	US-08-612-521-3	Sequence 3, Appli
	22	809	31.2	3792	4	US-09-351-224E-10	Sequence 10, Appl
	23	770	29.7	3924	2	US-08-996-644-3	Sequence 3, Appli
	24	770	29.7	3924	3	US-09-352-552-3	Sequence 3, Appli
	25	770	29.7	3927	2	US-08-996-644-1	Sequence 1, Appli
	26	770	29.7	3927	3	US-09-352-552-1	Sequence 1, Appli
	27	712	27.5	3999	4	US-09-351-224E-9	Sequence 9, Appli
	28	669.5	25.8	2376	1	US-08-394-880B-1	Sequence 1, Appli
	29	641	24.7	2061	4	US-09-061-764A-17	Sequence 17, Appl
	30	639.5	24.7	3909	1	US-08-232-537-1	Sequence 1, Appli
	31	635	24.5	1959	4	US-09-061-764A-4	Sequence 4, Appli
	32	598	23.1	2244	4	US-09-061-764A-18	Sequence 18, Appl
	33	581	22.4	1749	4	US-09-134-001C-1893	Sequence 1893, Ap
	34	543.5	21.0	1743	4	US-09-134-001C-2774	Sequence 2774, Ap
	35	529	20.4	7760	4	US-08-961-527-63	Sequence 63, Appl
	36	521	20.1	5120	3	US-08-772-270A-6	Sequence 6, Appli
	37	521	20.1	8370	2	US-08-488-706-1	Sequence 1, Appli
	38	498.5	19.2	13188	4	US-08-961-527-70	Sequence 70, Appl
	39	498.5	19.2	4411529	4	US-09-103-840A-1	Sequence 1, Appli
	40	497.5	19.2	7721	3	US-08-772-270A-14	Sequence 14, Appl
	41	496.5	19.2	4403765	4	US-09-103-840A-2	Sequence 2, Appli
c	42	495.5	19.1	28804	2	US-08-592-874-1	Sequence 1, Appli
c	43	495.5	19.1	28804	3	US-09-096-942-2	Sequence 2, Appli
c	44	495.5	19.1	28804	3	US-09-096-867-2	Sequence 2, Appli
	45	472.5	18.2	2407	2	US-08-895-522-2	Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-09-120-513-1
; Sequence 1, Application US/09120513
; Patent No. 6025160
;  GENERAL INFORMATION:
;    APPLICANT:  Brun, Kimberly
;    APPLICANT:  Chenery, Richard
;    APPLICANT:  Ellens, Harma
;    APPLICANT:  Field, John
;    APPLICANT:  Yue, Lin
;    TITLE OF INVENTION:  POLYNUCLEOTIDE AND POLYPEPTIDE
;    TITLE OF INVENTION:  SEQUENCES ENCODING RAT MDR1B2 AND
;    TITLE OF INVENTION:  SCREENING METHODS THEREOF
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corporation
;      STREET:  709 Swedeland Road
;      CITY:  King of Prussia
;      STATE:  PA
;      COUNTRY:
;      ZIP:  19406
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/120,513
;      FILING DATE:  22-JUL-1998
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  King, William T
;      REGISTRATION NUMBER:  30,954
;      REFERENCE/DOCKET NUMBER:  GP50008
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-5015
```

```
;      TELEFAX:  610-270-5090
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4233 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  Genomic DNA
US-09-120-513-1

Alignment Scores:
Pred. No.:              2.88e-179        Length:          4233
Score:                  1576.00          Matches:         308
Percent Similarity:     75.90%           Conservative:    92
Best Local Similarity:  58.44%           Mismatches:      105
Query Match:            60.80%           Indels:          22
DB:                     3                Gaps:            4

US-09-873-409-8 (1-514) x US-09-120-513-1 (1-4233)

Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
             |||   |||||    |||||::::::::::::|||        |||||||||:::    |||
Db      170 ATGGCTCTGGGAACTCTCGCTGCTATCATCCACGGAACCCTGCTTCCCCTCCTGATGCTG 229

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
             |||   |||   |||:::|||:::
Db      230 GTGTTCGGATACATGACAGATAGTTTACCCAAGCAGAGACCCGCATTCTGCCGAGCGTT 289

Qy       31 ------SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu--------- 45
                  :::          |||    |||    |||          |||
Db      290 ACTAATCAAAGTGAAATCAACAGTACACAGACCGTCAGCGACAGCAGTCTGGAGGAGGAC 349

Qy       46 ------------ThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                           ||||||   ||||||||          ||||||         |||
Db      350 ATGGCCATGTACGCCTACTATTACACGGGCATTGGTGCCGGTGTGCTCATCGTTGCCTAC 409

Qy       62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
            |||||:::||||||||    :::    |||    |||||          :::||||||::::::
Db      410 ATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAAG 469

Qy       82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
            ||||||||:::::::::    |||:::||||||||||||||           |||    ||||||
Db      470 TTTTTCCATGCCATCATGAATCAGGAGATAGGCTGGTTTGACGTGAATGACGCTGGGGAG 529

Qy      102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
            |||||||||||:::|||      |||:::    ||||||:::|||||||||||||||||||:::
Db      530 CTCAACACCCGGCTCACAGATGACGTCTCCAAAATTAATGACGGAATTGGTGACAAACTT 589

Qy      121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
             :::     ||||||:::::::::||||||||      |||          :::|||       :::
Db      590 GGAATGTTCTTTCAGTCCATAACGACATTTTCAGCCGGTTTTATAATAGGATTTATAAGT 649

Qy      141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
            ||||||||||||||||||     |||:::     ||||||||||||          |||:::|||
Db      650 GGTTGGAAGCTAACCCTTGTAATTTTGGCCGTCAGCCCTCTTATTGGGTTGTCATCTGCC 709

Qy      161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
                 ::::::::::::     |||    |||:::||||||||    ||||||:::||||||
Db      710 ATGTGGGCAAAGGTACTGACTTCATTTACTAATAAGGAACTCCAGGCTTATGCGAAAGCT 769

Qy      181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
            ||||||||||||||||||||||:::::::||||||||||||||||||          |||
Db      770 GGAGCAGTTGCCGAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCGTTTGGAGGACAA 829

Qy      201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
            :::|||||||||:::|||        :::||||||::::::||||||          |||||||||
Db      830 AAGAAGGAACTTGAAAGGTACAATAAAAATTTAGAAGAAGCTAAAAGAGTTGGCATAAAG 889

Qy      221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
            :::      |||    :::    :::|||:::|||          |||          :::    :::|||
Db      890 AAAGCCATCACGGCCAACATTTCCATAGGTATTGCCTACCTGTTGGTCTATGCGTCTTAT 949

Qy      241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
                |||||||||||||||||||||||||||:::|||:::     |||        |||:::|||
Db      950 GCACTGGCATTCTGGTATGGGACCTCCTTGGTCCTCTCAAATGAA------TATTCTATT 1003

Qy      261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
                |||   ||||||   ||||||||||||::::::        ::::::    ||||||
Db     1004 GGACAAGTGCTTACCGTCTTCTTCTCTATTTTATTGGGGACTTTCAGTATTGGACATTTA 1063

Qy      281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
                |||:::    |||    ||||||    ||||||||||||||||:::    ||||||::::::
Db     1064 GCCCCAAACATAGAAGCCTTTGCAAATGCAAGAGGGGCAGCCTATGAAATCTTCAAGATA 1123

Qy      301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
            ||||||    :::||||||||||||:::||||||||||    |||:::||||||:::||||||
Db     1124 ATTGATAATGAGCCAAGCATTGACAGCTTCTCAACCAAGGGACACAAACCAGACAGTATA 1183

Qy      321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
                |||    :::|||||||||||||||    |||||||||||||||||          :::|||
Db     1184 ATGGGAAATTTGGAATTTAAAAATGTTTACTTCAACTACCCATCACGAAGTGAAGTTAAG 1243

Qy      341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
            ||||||||||||||||||||:::::::||||||||||:::|||||||||||||||||||
Db     1244 ATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTGGTTGGCAAC 1303

Qy      361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
            :::|||    ||||||||||||    |||||||||||||||||||||||||||    :::
Db     1304 AGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCATAGAG 1363

Qy      381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
            |||    :::    :::|||          ||||||||    :::|||||||||:::    |||:::
Db     1364 GGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAA 1423

Qy      401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
                ||||||||||||||||||||||||||||||    ||||||||:::     ||||||
Db     1424 ATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCCACCACGATTGCCGAAAACATT 1483

Qy      421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
            :::||||||||:::::::::||||||    :::||||:::||||||    :::|||||||||
Db     1484 CGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAAT 1543

Qy      441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
            ||||||||||||||||||:::    |||:::||||||||||||||||||||||||:::|||
Db     1544 GCCTATGACTTCATCATGAAACTGCCCCACAAATTTAACACCCTGGTTGGTGAGAGAGGG 1603

Qy      461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
            ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     1604 GCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAAC 1663

Qy      481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
            ||||||||||||:::|||||||||||||||||||||||||||:::||||||::::::
Db     1664 CCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTG 1723

Qy      501 ValGlnAlaAlaLeuGluLys 507
            ||||||||||||||:::|||
Db     1724 GTTCAGGCCGCTCTGGATAAG 1744

RESULT 2
US-09-450-105-1
; Sequence 1, Application US/09450105
; Patent No. 6169166
; GENERAL INFORMATION:
;  APPLICANT: Kimberly Anne Brun
;  APPLICANT: Richard James Chenery
;  APPLICANT: Harma Ellens
;  APPLICANT: John Anthony Feild
;  APPLICANT: Lin Yue
;  TITLE OF INVENTION: POLYNUCLEOTIDE AND POLYPEPTIDE SEQUENCES
;  TITLE OF INVENTION: ENCODING RAT MDR1B2 AND SCREENING METHODS THEREOF
;  FILE REFERENCE: GP-50008-D1
;  CURRENT APPLICATION NUMBER: US/09/450,105
;  CURRENT FILING DATE: 1999-11-29
;  EARLIER APPLICATION NUMBER: 09/120,513
;  EARLIER FILING DATE: 1998-07-22
```

```
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 1
;   LENGTH: 4233
;   TYPE: DNA
;   ORGANISM: HOMO SAPIENS
US-09-450-105-1

Alignment Scores:
Pred. No.:              2.88e-179     Length:        4233
Score:                  1576.00       Matches:       308
Percent Similarity:     75.90%        Conservative:  92
Best Local Similarity:  58.44%        Mismatches:    105
Query Match:            60.80%        Indels:        22
DB:                     4             Gaps:          4

US-09-873-409-8 (1-514) x US-09-450-105-1 (1-4233)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
             |||   ||||||   ||||||:::::::::::|||        |||||||||:::    |||
Db     170 ATGGCTCTGGGAACTCTCGCTGCTATCATCCACGGAACCCTGCTTCCCCTCCTGATGCTG 229

Qy      21 ValLeuGlyGluMetSerAspAsnLeuIle------------------------------ 30
             |||    |||    |||:::|||:::
Db     230 GTGTTCGGATACATGACAGATAGTTTTACCCAAGCAGAGACCCGCATTCTGCCGAGCGTT 289

Qy      31 ------SerGlyCysLeuValGlnThrAsnThrTyrSerPhePheArgLeu--------- 45
                 :::         |||   |||   |||        |||
Db     290 ACTAATCAAAGTGAAATCAACAGTACACAGACCGTCAGCGACAGCAGTCTGGAGGAGGAC 349

Qy      46 ------------ThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                       ||||||   |||||||||            ||||||        |||
Db     350 ATGGCCATGTACGCCTACTATTACACGGGCATTGGTGCCGGTGTGCTCATCGTTGCCTAC 409

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
             ||||||:::|||||||||    :::    |||   ||||||        :::||||||::::::
Db     410 ATCCAGGTTTCACTTTGGTGCCTGGCAGCTGGGAGACAAATACACAAGATTAGGCAGAAG 469

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
             ||||||||:::::::::   |||:::|||||||||||||||        |||    ||||||
Db     470 TTTTTCCATGCCATCATGAATCAGGAGATAGGCTGGTTTGACGTGAATGACGCTGGGGAG 529

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
             ||||||||||||:::|||    |||:::   ||||||:::||||||||||||||||||:::
Db     530 CTCAACACCCGGCTCACAGATGACGTCTCCAAAATTAATGACGGAATTGGTGACAAACTT 589

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPhePheSerIleGlyLeuAlaValGlyLeuValLys 140
             :::      ||||||:::::::::|||||||||        :::|||    |||
Db     590 GGAATGTTCTTTCAGTCCATAACGACATTTTCAGCCGGTTTTATAATAGGATTTATAAGT 649

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
             |||||||||||||||||||||   |||:::   ||||||||||||        |||:::|||
Db     650 GGTTGGAAGCTAACCCTTGTAATTTTGGCCGTCAGCCCTCTTATTGGGTTGTCATCTGCC 709

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
                :::::::::::    |||    |||:::||||||||||    ||||||:::||||||
Db     710 ATGTGGGCAAAGGTACTGACTTCATTTACTAATAAGGAACTCCAGGCTTATGCGAAAGCT 769

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
             ||||||||||||||||||||||||::::::||||||||||||||||||||          |||
Db     770 GGAGCAGTTGCCGAAGAAGTCTTAGCAGCCATCAGAACTGTGATTGCGTTTGGAGGACAA 829

Qy     201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
             :::|||||||||:::||||||    :::||||||::::::||||||     |||||||||
Db     830 AAGAAGGAACTTGAAAGGTACAATAAAAATTTAGAAGAAGCTAAAAGAGTTGGCATAAAG 889

Qy     221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
             :::    |||    :::     :::|||:::|||          |||       :::      :::|||
Db     890 AAAGCCATCACGGCCAACATTTCCATAGGTATTGCCTACCTGTTGGTCTATGCGTCTTAT 949

Qy     241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
                ||||||||||||||||||||||||||||:::|||:::    |||       |||:::|||
Db     950 GCACTGGCATTCTGGTATGGGACCTCCTTGGTCCTCTCAAATGAA------TATTCTATT 1003

Qy     261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
             |||    ||||||    |||||||||||||::::::       ::::::    ||||||
Db    1004 GGACAAGTGCTTACCGTCTTCTTCTCTATTTATTGGGGACTTTCAGTATTGGACATTTA 1063

Qy     281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
                |||:::    |||    ||||||    ||||||||||||||||:::    ||||||::::::
Db    1064 GCCCCAAACATAGAAGCCTTTGCAAATGCAAGAGGGGCAGCCTATGAAATCTTCAAGATA 1123

Qy     301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
             ||||||    :::||||||||||||:::|||||||||    |||:::||||||:::||||||
Db    1124 ATTGATAATGAGCCAAGCATTGACAGCTTCTCAACCAAGGGACACAAACCAGACAGTATA 1183

Qy     321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
                |||    :::|||||||||||||||    ||||||||||||||||||          :::|||
Db    1184 ATGGGAAATTTGGAATTTAAAAATGTTTACTTCAACTACCCATCACGAAGTGAAGTTAAG 1243

Qy     341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
             |||||||||||||||||||::::::|||||||||:::||||||||||||||||||
Db    1244 ATCTTGAAGGGCCTCAACCTGAAGGTGAAGAGCGGGCAGACGGTAGCCCTGGTTGGCAAC 1303

Qy     361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
             :::|||    ||||||||||||    |||||||||||||||||||||||||||||    :::
Db    1304 AGTGGCTGTGGGAAAAGCACAACTGTCCAGCTGCTGCAGAGGCTCTACGACCCCATAGAG 1363

Qy     381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
             |||    :::    :::|||        ||||||||||    :::|||||||||:::    |||:::
Db    1364 GGCGAGGTCAGTATCGACGGACAGGACATCAGGACCATCAATGTGAGGTATCTGCGGGAA 1423

Qy     401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
             ||||||||||||||||||||||||||||||||    |||||||||:::     ||||||
Db    1424 ATCATTGGGGTGGTGAGTCAGGAACCCGTGCTGTTTGCCACCACGATTGCCGAAAACATT 1483

Qy     421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
             :::||||||||||||:::::::||||||     :::|||:::|||:::|||    :::|||||||||
Db    1484 CGCTATGGCCGAGAAAACGTCACCATGGATGAGATAGAGAAAGCTGTCAAGGAAGCCAAT 1543

Qy     441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
             ||||||||||||||||||:::    |||:::|||||||||||||||||||||||:::|||
Db    1544 GCCTATGACTTCATCATGAAACTGCCCCACAAATTTAACACCCTGGTTGGTGAGAGAGGG 1603

Qy     461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
             ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1604 GCGCAGCTGAGTGGGGGACAGAAACAGAGGATCGCCATTGCCCGGGCCCTGGTCCGCAAC 1663

Qy     481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
             ||||||||||||:::||||||||||||||||||||||||||:::||||||::::::
Db    1664 CCCAAGATCCTTTTGTTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCCGTG 1723

Qy     501 ValGlnAlaAlaLeuGluLys 507
             |||||||||||||||:::|||
Db    1724 GTTCAGGCCGCTCTGGATAAG 1744

RESULT 3
US-08-752-447-1
; Sequence 1, Application US/08752447
; Patent No. 5994088
;  GENERAL INFORMATION:
;    APPLICANT:  Mechetner, Eugene
;    APPLICANT:  Roninson, Igor B
;    TITLE OF INVENTION:  Methods and Reagents for Preparing and
;    TITLE OF INVENTION:  Using Immunoligcal Agents Specific for P-glycoprotein
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McDonnell Boehnen Hulbert & Berghoff Ltd.
;      STREET:  300 South Wacker Drive, Seventh Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/752,447
;      FILING DATE:  15-NOV-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5994088nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  95,1121
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-913-0001
;      TELEFAX:  312-913-9808
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4669 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  5'UTR
;      LOCATION:  1..424
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  425..4264
;    FEATURE:
;      NAME/KEY:  3'UTR
;      LOCATION:  4265..4669
US-08-752-447-1

Alignment Scores:
Pred. No.:              3.4e-179      Length:          4669
Score:                  1576.00       Matches:         305
Percent Similarity:     77.04%        Conservative:    101
Best Local Similarity:  57.87%        Mismatches:      99
Query Match:            60.80%        Indels:          22
DB:                     2             Gaps:            5

US-09-873-409-8 (1-514) x US-08-752-447-1 (1-4669)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy     21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db    635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe----------------- 43
Db    695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    875 TTTTTTCATGCTATTATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    935 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAAATT 994

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db   1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db   1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db   1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db   1235 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1294

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db   1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1354

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db   1355 GCTGTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1408

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db   1409 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1468

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db   1469 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db   1529 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1588

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db   1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1648

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   1649 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1708

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   1709 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1768

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   1769 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1828

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   1829 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1888

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1889 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1948

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1949 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2008

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   2009 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2068

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   2069 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2128

Qy    501 ValGlnAlaAlaLeuGluLys 507
```

```
             ||||||   ||||||:::|||
Db    2129 GTTCAGGTGGCTCTGGATAAG 2149

RESULT 4
US-09-316-167-1
; Sequence 1, Application US/09316167
; Patent No. 6365357
;  GENERAL INFORMATION:
;    APPLICANT:  Mechetner, Eugene
;    APPLICANT:  Roninson, Igor B
;    TITLE OF INVENTION:  Methods and Reagents for Preparing and
;    TITLE OF INVENTION:  Using Immunoligcal Agents Specific for P-glycoprotein
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  McDonnell Boehnen Hulbert & Berghoff Ltd.
;      STREET:  300 South Wacker Drive, Seventh Floor
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  USA
;      ZIP:  60606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/316,167
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/752,447
;      FILING DATE:  15-NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No.'6365357nan, Kevin E
;      REGISTRATION NUMBER:  35,303
;      REFERENCE/DOCKET NUMBER:  95,1121
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  312-913-0001
;      TELEFAX:  312-913-9808
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4669 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    FEATURE:
;      NAME/KEY:  5'UTR
;      LOCATION:  1..424
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  425..4264
;    FEATURE:
;      NAME/KEY:  3'UTR
;      LOCATION:  4265..4669
US-09-316-167-1

Alignment Scores:
Pred. No.:                3.4e-179     Length:          4669
Score:                    1576.00      Matches:         305
Percent Similarity:       77.04%       Conservative:    101
Best Local Similarity:    57.87%       Mismatches:      99
Query Match:              60.80%       Indels:          22
DB:                       4            Gaps:            5

US-09-873-409-8 (1-514) x US-09-316-167-1 (1-4669)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
             |||:::::|||   |||||:::::::::::::||||||   ||||||||||||    |||
Db     575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy      21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
             |||   |||||||||:::|||                           :::|||:::|||
Db     635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                     ::::::      |||         ||||||
Db     695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                 |||         ||||||  |||||||||        |||:::       |||
Db     755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
           ||||||:::|||   |||   :::   |||   ||||||           :::||||||||||||
Db     815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
           |||||||||:::::::::   |||:::||||||||||||||||       |||:::||||||
Db     875 TTTTTCATGCTATTATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
           ||||||||||||:::|||   |||:::   ||||||:::::|||||||||||||||||||
Db     935 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAAATT 994

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
              :::   ||||||:::|||:::||||||      |||      ||||||      :::
Db     995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
           ||||||||||||||||||||||||   |||:::   ||||||:::::::      |||||||||
Db    1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
                 ::::::::::::      |||   |||   |||||||||   ||||||:::||||||
Db    1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
           ||||||||||||||||||||||||:::::::|||||||||||||||||||      |||
Db    1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy     201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
           :::|||||||||:::||||||   :::||||||:::::::||||||      |||||||||
Db    1235 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1294

Qy     221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
           :::   |||   :::   :::|||:::||||||      :::         :::|||
Db    1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1354

Qy     241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
              :::||||||||||||||||||:::|||:::|||:::||||||      ||||:::|||
Db    1355 GCTGTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1408

Qy     261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
           |||   ||||||   |||||||||||||||:::      ::::::   :::|||   |||
Db    1409 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1468

Qy     281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
              |||      |||   ||||||   |||||||||||||||:::   ||||||::::::
Db    1469 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528

Qy     301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
           ||||||   |||||||||||||||:::::::|||   :::|||:::|||||||:::::::|||
Db    1529 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1588

Qy     321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
           :::|||   :::||||||:::||||||   |||:::|||||||||||||      :::|||
Db    1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1648

Qy     341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
           |||||||||||||||||||:::::::::||||||:::||||||||||||||||||
Db    1649 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1708

Qy     361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
           :::|||   ||||||||||||   |||||||||:::|||||||||||||||||||   :::
Db    1709 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1768
```

```
Qy     381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
           |||   :::   |||||      ||||||||   :::||||||||     |||:::
Db    1769 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1828

Qy     401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
              |||||||||||||||||||||||||||||||||    |||||||||:::   ||||||
Db    1829 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1888

Qy     421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
           :::|||||||||::::::||||||   :::|||:::|||:::|||   :::|||||||||
Db    1889 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1948

Qy     441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
           ||||||||||||||||||:::   |||:::||||||:::||||||||||||||:::|||
Db    1949 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2008

Qy     461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
           ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||
Db    2009 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2068

Qy     481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
           |||||||||||||:::||||||||||||||||||||||||||||:::||||||::::::
Db    2069 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2128

Qy     501 ValGlnAlaAlaLeuGluLys 507
           ||||||   ||||||:::|||
Db    2129 GTTCAGGTGGCTCTGGATAAG 2149
```

RESULT 5
US-08-181-471-2
; Sequence 2, Application US/08181471
; Patent No. 5641508
; GENERAL INFORMATION:
;   APPLICANT: Li, Lingna
;   APPLICANT: Lishko, Valeryi K.
;   TITLE OF INVENTION: METHOD FOR DELIVERING BENEFICIAL
;   TITLE OF INVENTION: COMPOSITIONS TO HAIR FOLLICLES
;   NUMBER OF SEQUENCES: 3
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Thomas Fitting
;   STREET: 12526 High Bluff Drive, Suite 300
;   CITY: San Diego
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 92130
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/181,471
;   FILING DATE: 13-JAN-1994
;   CLASSIFICATION: 424
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/041,553
;   FILING DATE: 02-APR-1993
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Fitting, Thomas
;   REGISTRATION NUMBER: 34,163
;   REFERENCE/DOCKET NUMBER: ANT0029P
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 619-792-3680
;   TELEFAX: 619-792-8477
; INFORMATION FOR SEQ ID NO: 2:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4646 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: cDNA
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 425..4267
US-08-181-471-2

```
Alignment Scores:
Pred. No.:              1.79e-178    Length:         4646
Score:                  1570.00      Matches:        305
Percent Similarity:     76.85%       Conservative:   100
Best Local Similarity:  57.87%       Mismatches:     100
Query Match:            60.57%       Indels:         22
DB:                     1            Gaps:           5
```

US-09-873-409-8 (1-514) x US-08-181-471-2 (1-4646)

```
Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||::::::|||   ||||||::::::::::::|||||   ||||||||||   |||
Db     575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy      21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
           |||   |||||||||:::|||                           :::|||:::|||
Db     635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                    ::::::      |||      ||||||
Db     695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                 |||         ||||||   |||||||||         |||:::      |||
Db     755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
           ||||||:::|||   |||   :::   |||   ||||||      :::|||||||||||||
Db     815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
           ||||||||::::::::::   |||:::||||||||||||||||      |||:::||||||
Db     875 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
           ||||||||||||:::|||   |||:::   ||||||::::::   ||||||||||||||||
Db     935 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 994

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
              :::   ||||||:::|||:::||||||      |||      ||||||      :::
Db     995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
           ||||||||||||||||||||||   |||:::   ||||||::::::      |||||||||
Db    1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
              ::::::::::::   |||   |||   |||||||||      ||||||:::||||||
Db    1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
           ||||||||||||||||||||||||::::::||||||||||||||||||||||      |||
Db    1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy     201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
           :::||||||||||:::||||||   :::||||||::::::||||||      |||||||||
Db    1235 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1294

Qy     221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
           :::   |||   :::   :::|||:::||||||   :::         :::      :::|||
Db    1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1354

Qy     241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
           ||||||||||||||||||||||||:::|||:::|||:::||||||      |||:::|||
Db    1355 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1408

Qy     261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
```

```
Db   1409 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTAGTGTTGGACAGGCA 1468

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db   1469 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db   1529 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1588

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db   1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1648

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   1649 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1708

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   1709 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1768

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   1769 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1828

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   1829 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1888

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1889 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1948

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1949 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2008

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   2009 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2068

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   2069 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2128

Qy    501 ValGlnAlaAlaLeuGluLys 507
Db   2129 GTTCAGGTGGCTCTGGATAAG 2149
```

RESULT 6
US-08-793-610-5
; Sequence 5, Application US/08793610
; Patent No. 5858744
; GENERAL INFORMATION:
;   APPLICANT: BAUM, Christopher
;   APPLICANT: STOCKING-HARBERS, Carol
;   APPLICANT: OSTERTAG, Wolfram
;   TITLE OF INVENTION: RETROVIRAL VECTOR HYBRIDS AND THE USE THEROF
;   TITLE OF INVENTION: FOR GENE TRANSFER
;   NUMBER OF SEQUENCES: 6
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
;   STREET: 655 Fifteenth Street N.W. Suite 330
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: U.S.A.
;   ZIP: 20005-5701
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/793,610
;   FILING DATE: 07-MAR-1997
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: DE P 44 31 973.8
;   FILING DATE: 08-SEP-1994
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: DE 195 03 952.1
;   FILING DATE: 07-FEB-1995
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PCT/EP95/03175
;   FILING DATE: 10-AUG-1995
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Berman, Richard J.
;   REGISTRATION NUMBER: 39,105
;   REFERENCE/DOCKET NUMBER: P1614-7007
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (202)638-5000
;   TELEFAX: (202)638-4810
; INFORMATION FOR SEQ ID NO: 5:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 6505 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: circular
;   MOLECULE TYPE: DNA
US-08-793-610-5

Alignment Scores:

Pred. No.:	3.17e-178	Length:	6505
Score:	1570.00	Matches:	305
Percent Similarity:	76.85%	Conservative:	100
Best Local Similarity:	57.87%	Mismatches:	100
Query Match:	60.57%	Indels:	22
DB:	2	Gaps:	5

US-09-873-409-8 (1-514) x US-08-793-610-5 (1-6505)

```
Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db   1967 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 2026

Qy     21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
Db   2027 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 2086

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db   2087 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 2146

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db   2147 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 2206

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db   2207 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 2266

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db   2267 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 2326

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db   2327 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 2386

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db   2387 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 2446

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
```

```
Db   2447 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 2506

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
              :::::::::::   |||   |||   ||||||||   |||||:::||||||
Db   2507 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 2566

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
          ||||||||||||||||||||||:::::||||||||||||||||||        |||
Db   2567 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 2626

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
          :::||||||||:::|||||    :::||||||::::::||||||      ||||||||
Db   2627 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 2686

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
          :::   |||   :::   :::||||:::||||||   :::      :::      :::|||
Db   2687 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 2746

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
             |||||||||||||||||||:::|||:::|||:::||||||      |||:::|||
Db   2747 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 2800

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
          |||   ||||||   |||||||||||||||:::      ::::::   :::|||   |||
Db   2801 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTAGTGTTGGACAGGCA 2860

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
             |||      |||   ||||||   |||||||||||||||:::   ||||||::::::
Db   2861 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 2920

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
          ||||||   |||||||||||||||::::::|||   :::|||:::||||||::::::|||
Db   2921 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 2980

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
          :::|||   :::||||||:::||||||   |||:::||||||||||||      :::|||
Db   2981 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 3040

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
          ||||||||||||||||||||||:::::::::||||||:::||||||||||||||||||
Db   3041 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 3100

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
          :::|||   ||||||||||||   ||||||||||:::||||||||||||||||||   :::
Db   3101 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 3160

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
          |||   :::   ||||||      |||||||||||   :::||||||      |||:::
Db   3161 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 3220

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
             ||||||||||||||||||||||||||||||||   ||||||||:::   ||||||
Db   3221 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 3280

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
          :::|||||||||::::::||||||   :::|||:::|||:::|||   :::|||||||||
Db   3281 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 3340

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
          ||||||||||||||||||:::   |||:::||||||:::||||||||||||||:::|||
Db   3341 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 3400

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
          ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||
Db   3401 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 3460

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
          ||||||||||||:::|||||||||||||||||||||||||||:::||||||::::::
Db   3461 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 3520

Qy    501 ValGlnAlaAlaLeuGluLys 507
          ||||||   ||||||:::|||
Db   3521 GTTCAGGTGGCTCTGGATAAG 3541
```

```
RESULT 7
US-08-793-610-6
; Sequence 6, Application US/08793610
; Patent No. 5858744
;  GENERAL INFORMATION:
;    APPLICANT:  BAUM, Christopher
;    APPLICANT:  STOCKING-HARBERS, Carol
;    APPLICANT:  OSTERTAG, Wolfram
;    TITLE OF INVENTION:  RETROVIRAL VECTOR HYBRIDS AND THE USE THEROF
;    TITLE OF INVENTION:  FOR GENE TRANSFER
;    NUMBER OF SEQUENCES:  6
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Nikaido, Marmelstein, Murray & Oram LLP
;      STREET:  655 Fifteenth Street N.W. Suite 330
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  U.S.A.
;      ZIP:  20005-5701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/793,610
;      FILING DATE:  07-MAR-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE P 44 31 973.8
;      FILING DATE:  08-SEP-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  DE 195 03 952.1
;      FILING DATE:  07-FEB-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/EP95/03175
;      FILING DATE:  10-AUG-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Berman, Richard J.
;      REGISTRATION NUMBER:  39,105
;      REFERENCE/DOCKET NUMBER:  P1614-7007
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (202)638-5000
;      TELEFAX:  (202)638-4810
;  INFORMATION FOR SEQ ID NO:  6:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  9318 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  circular
;    MOLECULE TYPE:  DNA
US-08-793-610-6

Alignment Scores:
Pred. No.:              5.85e-178     Length:         9318
Score:                  1570.00       Matches:        305
Percent Similarity:     76.85%        Conservative:   100
Best Local Similarity:  57.87%        Mismatches:     100
Query Match:            60.57%        Indels:         22
DB:                     2             Gaps:           5

US-09-873-409-8 (1-514) x US-08-793-610-6 (1-9318)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
          |||:::::|||   ||||||::::::::::::||||||   ||||||||||||   |||
Db   1926 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 1985

Qy     21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
          |||   |||||||||:::|||                          :::|||:::|||
Db   1986 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 2045

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                          ::::::   |||         ||||||
Db   2046 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 2105
```

```
Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db   2106 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 2165

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db   2166 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 2225

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db   2226 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 2285

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db   2286 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAAGTTATTGGTGACAAAATT 2345

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db   2346 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 2405

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db   2406 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 2465

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db   2466 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 2525

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db   2526 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 2585

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db   2586 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 2645

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db   2646 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 2705

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db   2706 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 2759

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db   2760 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 2819

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db   2820 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 2879

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db   2880 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 2939

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db   2940 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 2999

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   3000 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 3059

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   3060 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 3119

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   3120 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 3179

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   3180 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 3239

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   3240 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 3299

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   3300 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 3359

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   3360 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 3419

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   3420 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 3479

Qy    501 ValGlnAlaAlaLeuGluLys 507
Db   3480 GTTCAGGTGGCTCTGGATAAG 3500
```

```
RESULT 8
US-08-784-649A-1
; Sequence 1, Application US/08784649A
; Patent No. 5830697
;  GENERAL INFORMATION:
;    APPLICANT:  Sikic, Branimir I
;    APPLICANT:  Chen, Gang
;    TITLE OF INVENTION:  P-GLYCOPROTEIN MUTANT RESISTANT TO
;    TITLE OF INVENTION:  CYCLOSPORIN MODULATION
;    NUMBER OF SEQUENCES:  5
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Fish & Richardson
;      STREET:  2200 Sand Hill Road
;      CITY:  Menlo Park
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94025
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/784,649A
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Sherwood, Pamela J
;      REGISTRATION NUMBER:  Reg.No. 5830697 36,677
;      REFERENCE/DOCKET NUMBER:  06037/007001
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415-322-5070
;      TELEFAX:  415-854-0875
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4264 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-784-649A-1

Alignment Scores:
Pred. No.:              2.34e-178     Length:         4264
Score:                  1568.50       Matches:        306
Percent Similarity:     77.04%        Conservative:   100
Best Local Similarity:  58.06%        Mismatches:     98
Query Match:            60.51%        Indels:         23
DB:                     2             Gaps:           6
```

```
US-09-873-409-8 (1-514) x US-08-784-649A-1 (1-4264)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
           |||::::::|||  ||||||:::::::::::::||||||  ||||||||||||  |||
Db     292 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 351

Qy      21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
           |||  ||||||||:::|||                          :::|||:::|||
Db     352 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 411

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe----------------- 43
                    ::::::      |||       ||||||
Db     412 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 471

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                 |||        ||||||  |||||||||        |||:::       |||
Db     472 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 531

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
           ||||||:::|||   |||   :::   |||   ||||||        :::||||||||||||
Db     532 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 591

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
           |||||||||:::::::::   |||:::|||||||||||||||||       |||:::||||||
Db     592 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 651

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
           ||||||||||||:::|||   :::|||:::   ||||||||||||||||||||||||||||||
Db     652 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAAATT 711

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
                 :::||||||:::|||:::||||||        |||        ||||||      :::
Db     712 GGA---ATGTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 768

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
           ||||||||||||||||||||||||   |||:::   ||||||::::::      |||||||||
Db     769 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 828

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
                 ::::::::::::   |||   |||   |||||||||   ||||||:::||||||
Db     829 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 888

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
           |||||||||||||||||||||||||:::::|||||||||||||||||||||      |||
Db     889 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 948

Qy     201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
           :::||||||||:::||||||   :::||||||::::::||||||      ||||||||||
Db     949 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1008

Qy     221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
           :::   |||   :::   :::|||:::||||||   :::      :::      :::|||
Db    1009 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1068

Qy     241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
              ||||||||||||||||||||:::|||:::|||:::||||||      |||:::|||
Db    1069 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1122

Qy     261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
           |||   ||||||   ||||||||||||||||:::      ::::::   :::|||   |||
Db    1123 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1182

Qy     281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
              |||      |||   ||||||   ||||||||||||||:::   ||||||::::::
Db    1183 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1242

Qy     301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
           ||||||   ||||||||||||||||::::::|||   :::|||:::||||||::::::|||
Db    1243 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1302

Qy     321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
           :::|||   :::||||||:::||||||   |||:::||||||||||||      :::|||
Db    1303 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1362

Qy     341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
           |||||||||||||||||||||:::::::::||||||:::|||||||||||||||||
Db    1363 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1422

Qy     361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
           :::|||   ||||||||||||   |||||||||:::|||||||||||||||||   :::
Db    1423 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1482

Qy     381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
           |||   :::   ||||||      |||||||||   :::|||||||||      |||:::
Db    1483 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1542

Qy     401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
              |||||||||||||||||||||||||||||   |||||||||:::   ||||||
Db    1543 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1602

Qy     421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
           :::||||||||||::::::||||||   :::|||:::|||:::|||   :::|||||||||
Db    1603 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1662

Qy     441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
           |||||||||||||||||:::   |||:::||||||:::||||||||||||||||:::|||
Db    1663 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1722

Qy     461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
           ||||||:::|||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1723 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1782

Qy     481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
           ||||||||||:::|||||||||||||||||||||||||||||:::||||||::::::
Db    1783 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1842

Qy     501 ValGlnAlaAlaLeuGluLys 507
           ||||||   ||||||:::|||
Db    1843 GTTCAGGTGGCTCTGGATAAG 1863
```

RESULT 9
US-08-784-649A-5
; Sequence 5, Application US/08784649A
; Patent No. 5830697
; GENERAL INFORMATION:
;   APPLICANT: Sikic, Branimir I
;   APPLICANT: Chen, Gang
;   TITLE OF INVENTION: P-GLYCOPROTEIN MUTANT RESISTANT TO
;   TITLE OF INVENTION: CYCLOSPORIN MODULATION
;   NUMBER OF SEQUENCES: 5
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Fish & Richardson
;   STREET: 2200 Sand Hill Road
;   CITY: Menlo Park
;   STATE: CA
;   COUNTRY: USA
;   ZIP: 94025
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/784,649A
;   FILING DATE:
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Sherwood, Pamela J
;   REGISTRATION NUMBER: Reg.No. 5830697 36,677
;   REFERENCE/DOCKET NUMBER: 06037/007001
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 415-322-5070
;   TELEFAX: 415-854-0875
; INFORMATION FOR SEQ ID NO: 5:
;   SEQUENCE CHARACTERISTICS:

```
;      LENGTH:  4264 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-784-649A-5

Alignment Scores:
Pred. No.:                 2.34e-178      Length:          4264
Score:                     1568.50        Matches:         306
Percent Similarity:        77.04%         Conservative:    100
Best Local Similarity:     58.06%         Mismatches:      98
Query Match:               60.51%         Indels:          23
DB:                        2              Gaps:            6

US-09-873-409-8 (1-514) x US-08-784-649A-5 (1-4264)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
             |||::::::|||   ||||||::::::::::::||||||  |||||||||||    |||
Db     292 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 351

Qy      21 ValLeuGlyGluMetSerAsp---------------------------AsnLeuIleSer 31
             |||   ||||||||:::|||                           :::|||:::|||
Db     352 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 411

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
                             ::::::      |||        ||||||
Db     412 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 471

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
                   |||        ||||||  ||||||||        |||:::        |||
Db     472 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 531

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
             ||||||:::|||   |||   :::   |||   ||||||      :::||||||||||||
Db     532 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 591

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
             ||||||||:::::::::   |||:::||||||||||||||      |||:::||||||
Db     592 TTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 651

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
             |||||||||||:::|||   |||:::   ||||||::::::|||||||||||||||||||
Db     652 CTTAACACCCGACTTACAGATGATGTCTCCAAGATTAATGAAGGAATTGGTGACAAAATT 711

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
                   :::||||||:::|||:::||||||      |||      ||||||      :::
Db     712 GGA---ATGTTCCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 768

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
             ||||||||||||||||||||||   |||:::      ||||||::::::   |||||||||
Db     769 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 828

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
                :::::::::::   |||   |||   |||||||||      ||||||:::||||||
Db     829 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 888

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
             |||||||||||||||||||||||||::::::|||||||||||||||||||      |||
Db     889 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 948

Qy     201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
             :::|||||||||:::||||||   :::||||||::::::||||||      |||||||||
Db     949 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1008

Qy     221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
             :::   |||   :::      :::|||:::||||||     :::       :::   :::|||
Db    1009 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1068

Qy     241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
                |||||||||||||||||||||:::|||:::|||:::||||||      |||:::|||
Db    1069 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1122

Qy     261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
             |||   ||||||   ||||||||||||||||:::          ::::::   :::|||   |||
Db    1123 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1182

Qy     281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
                |||      |||   ||||||   |||||||||||||||:::   ||||||::::::
Db    1183 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1242

Qy     301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
             ||||||   ||||||||||||||||::::::|||   :::|||:::||||||::::::|||
Db    1243 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1302

Qy     321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
             :::|||   :::||||||:::||||||   |||:::|||||||||||      :::|||
Db    1303 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1362

Qy     341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
             ||||||||||||||||||||::::::::::||||||:::||||||||||||||||||
Db    1363 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1422

Qy     361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
             :::|||   |||||||||||   ||||||||:::|||||||||||||||||   :::
Db    1423 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1482

Qy     381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
             |||   :::   ||||||      |||||||||   :::||||||||||      |||:::
Db    1483 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1542

Qy     401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
                ||||||||||||||||||||||||||||||   |||||||||:::   ||||||
Db    1543 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1602

Qy     421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
             :::||||||||:::||||||   :::||:::||:::|||   :::|||||||||
Db    1603 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1662

Qy     441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
             ||||||||||||||||||:::   |||:::||||||:::||||||||||||||||:::|||
Db    1663 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 1722

Qy     461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
             ||||||:::||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    1723 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 1782

Qy     481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
             ||||||||||:::||||||||||||||||||||||||||:::||||||::::::
Db    1783 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 1842

Qy     501 ValGlnAlaAlaLeuGluLys 507
             ||||||   ||||||:::|||
Db    1843 GTTCAGGTGGCTCTGGATAAG 1863

RESULT 10
5206352-3
;Patent No. 5206352
;    APPLICANT: Roninson, Igor B.;Pastan Ira H.;Gottesman,
;Michael M.
;    TITLE OF INVENTION: COMPOSITIONS FOR CLONES CONTAINING DNA
;SEQUENCES ASSOCIATED WITH MULTIDRUG RESISTANCE IN HUMAN CELLS
;    NUMBER OF SEQUENCES: 4
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/07/622,836
;      FILING DATE: 24-SEP-1990
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER: 892,575
;      FILING DATE: 01-AUG-1986
;      APPLICATION NUMBER: 845,610
;      FILING DATE: 28-MAR-1986
;SEQ ID NO:3:
;      LENGTH: 4669
5206352-3

Alignment Scores:
```

```
Pred. No.:              3.14e-178    Length:          4669
Score:                  1568.00      Matches:         305
Percent Similarity:     76.66%       Conservative:    99
Best Local Similarity:  57.87%       Mismatches:      101
Query Match:            60.49%       Indels:          22
DB:                     6            Gaps:            5

US-09-873-409-8 (1-514) x 5206352-3 (1-4669)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy     21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db    635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy     32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db    695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy     44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db    755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy     62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db    815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy     82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db    875 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGGAG 934

Qy    102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db    935 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATCAAGTTATTGGTGACAAAATT 994

Qy    121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db    995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy    141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db   1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy    161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db   1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy    181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db   1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy    201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db   1235 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1294

Qy    221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db   1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1354

Qy    241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db   1355 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1408

Qy    261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db   1409 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1468

Qy    281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db   1469 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528

Qy    301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db   1529 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1588

Qy    321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db   1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1648

Qy    341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db   1649 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTGGCCCTGGTTGGAAAC 1708

Qy    361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db   1709 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1768

Qy    381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db   1769 GGGATGGTCAGTGTTGATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1828

Qy    401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db   1829 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1888

Qy    421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db   1889 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1948

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db   1949 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2008

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db   2009 GCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2068

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db   2069 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2128

Qy    501 ValGlnAlaAlaLeuGluLys 507
Db   2129 GTTCAGGTGGCTCTGGATAAG 2149

RESULT 11
US-08-583-276-18
; Sequence 18, Application US/08583276
; Patent No. 5837536
;  GENERAL INFORMATION:
;    APPLICANT:  McDonagh, Kevin T.
;    APPLICANT:  Nienhuis, Arthur
;    APPLICANT:  Tolstoshev, Paul
;    TITLE OF INVENTION:  IMPROVED EXPRESSION OF HUMAN
;    TITLE OF INVENTION:  MULTIDRUG RESISTANCE GENES AND IMPROVED
;    TITLE OF INVENTION:  SELECTION OF CELLS TRANSDUCED WITH SUCH GENES
;    NUMBER OF SEQUENCES:  19
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Carella, Byrne, Bain, Gilfillan,
;      ADDRESSEE:  Cecchi & Stewart
;      STREET:  6 Becker Farm Road
;      CITY:  Roseland
;      STATE:  New Jersey
;      COUNTRY:  USA
;      ZIP:  07068
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch diskette
;      COMPUTER:  IBM PS/2
;      OPERATING SYSTEM:  PC-DOS
;      SOFTWARE:  DW4.V2
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/583,276
;      FILING DATE:  05-JAN-1996
;      CLASSIFICATION:  435
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/332,444
;      FILING DATE:  31-OCT-1994
;      APPLICATION NUMBER:  07/887,712
;      FILING DATE:  22-MAY-1992
;  INFORMATION FOR SEQ ID NO:  18:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4669 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  singular
;      TOPOLOGY:  linear
;    MOLECULE TYPE:
;      DESCRIPTION:  Genomic DNA
US-08-583-276-18

Alignment Scores:
Pred. No.:               6.15e-176      Length:          4669
Score:                   1549.00        Matches:         301
Percent Similarity:      76.28%         Conservative:    101
Best Local Similarity:   57.12%         Mismatches:      103
Query Match:             59.76%         Indels:          22
DB:                      2              Gaps:            5

US-09-873-409-8 (1-514) x US-08-583-276-18 (1-4669)

Qy       1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db     575 ATGGTGGTGGGAACTTTGGCTGCCATCATCCATGGGGCTGGACTTCCTCTCATGATGCTG 634

Qy      21 ValLeuGlyGluMetSerAsp--------------------------AsnLeuIleSer 31
Db     635 GTGTTTGGAGAAATGACAGATATCTTTGCAAATGCAGGAAATTTAGAAGATCTGATGTCA 694

Qy      32 GlyCysLeuValGlnThr------AsnThrTyrSerPhePhe------------------ 43
Db     695 AACATCACTAATAGAAGTGATATCAATGATACAGGGTTCTTCATGAATCTGGAGGAAGAC 754

Qy      44 ------ArgLeuThrLeuTyrTyrValGlyIleGlyValAlaAlaLeuIlePheGlyTyr 61
Db     755 ATGACCAGGTATGCCTATTATTACAGTGGAATTGGTGCTGGGGTGCTGGTTGCTGCTTAC 814

Qy      62 IleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIleArgLysGln 81
Db     815 ATTCAGGTTTCATTTTGGTGCCTGGCAGCTGGAAGACAAATACACAAAATTAGAAAACAG 874

Qy      82 PhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAspIleGlyGlu 101
Db     875 TTTTTTCATGCTATAATGCGACAGGAGATAGGCTGGTTTGATGTGCACGATGTTGGGCAG 934

Qy     102 LeuAsnThrArgMetThr---AspIleAspLysIleSerAspGlyIleGlyAspLysIle 120
Db     935 CTTAACACCCGACTTACAGATGATGTCTCTAAGATTAATGAACTTATTGGTGACAAAATT 994

Qy     121 AlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGlyLeuValLys 140
Db     995 GGAATGTTCTTTCAGTCAATGGCAACATTTTTCACTGGGTTTATAGTAGGATTTACACGT 1054

Qy     141 GlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAlaSerAlaAla 160
Db    1055 GGTTGGAAGCTAACCCTTGTGATTTTGGCCATCAGTCCTGTTCTTGGACTGTCAGCTGCT 1114

Qy     161 AlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyrSerLysAla 180
Db    1115 GTCTGGGCAAAGATACTATCTTCATTTACTGATAAAGAACTCTTAGCGTATGCAAAAGCT 1174

Qy     181 GlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPheArgAlaGln 200
Db    1175 GGAGCAGTAGCTGAAGAGGTCTTGGCAGCAATTAGAACTGTGATTGCATTTGGAGGACAA 1234

Qy     201 GluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLysAspPheGlyIleLys 220
Db    1235 AAGAAAGAACTTGAAAGGTACAACAAAAATTTAGAAGAAGCTAAAAGAATTGGGATAAAG 1294

Qy     221 ArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMetAsnGlyThrTyr 240
Db    1295 AAAGCTATTACAGCCAATATTTCTATAGGTGCTGCTTTCCTGCTGATCTATGCATCTTAT 1354

Qy     241 GlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluProGlyTyrThrIle 260
Db    1355 GCTCTGGCCTTCTGGTATGGGACCACCTTGGTCCTCTCAGGGGAA------TATTCTATT 1408

Qy     261 GlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCysIleGlyAlaAla 280
Db    1409 GGACAAGTACTCACTGTATTCTTTTCTGTATTAATTGGGGCTTTTAGTGTTGGACAGGCA 1468

Qy     281 ValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHisIlePheGlnVal 300
Db    1469 TCTCCAAGCATTGAAGCATTTGCAAATGCAAGAGGAGCAGCTTATGAAATCTTCAAGATA 1528

Qy     301 IleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLysProGluSerIle 320
Db    1529 ATTGATAATAAGCCAAGTATTGACAGCTATTCGAAGAGTGGGCACAAACCAGATAATATT 1588

Qy     321 GluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArgProSerIleLys 340
Db    1589 AAGGGAAATTTGGAATTCAGAAATGTTCACTTCAGTTACCCATCTCGAAAAGAAGTTAAG 1648

Qy     341 IleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAlaLeuValGlyLeu 360
Db    1649 ATCTTGAAGGGCCTGAACCTGAAGGTGCAGAGTGGGCAGACGGTCCCCCTGGTTGGAAAC 1708

Qy     361 AsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyrAspProAspAsp 380
Db    1709 AGTGGCTGTGGGAAGAGCACAACAGTCCAGCTGATGCAGAGGCTCTATGACCCCACAGAG 1768

Qy     381 GlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArgHisTyrArgAsp 400
Db    1769 GGGATGGTCAGTGTTCATGGACAGGATATTAGGACCATAAATGTAAGGTTTCTACGGGAA 1828

Qy     401 HisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIleSerAsnAsnIle 420
Db    1829 ATCATTGGTGTGGTGAGTCAGGAACCTGTATTGTTTGCCACCACGATAGCTGAAAACATT 1888

Qy     421 LysTyrGlyArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
Db    1889 CGCTATGGCCGTGAAAATGTCACCATGGATGAGATTGAGAAAGCTGTCAAGGAAGCCAAT 1948

Qy     441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
Db    1949 GCCTATGACTTTATCATGAAACTGCCTCATAAATTTGACACCCTGGTTGGAGAGAGAGGG 2008

Qy     461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
Db    2009 CCCCAGTTGAGTGGTGGGCAGAAGCAGAGGATCGCCATTGCACGTGCCCTGGTTCGCAAC 2068

Qy     481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
Db    2069 CCCAAGATCCTCCTGCTGGATGAGGCCACGTCAGCCTTGGACACAGAAAGCGAAGCAGTG 2128

Qy     501 ValGlnAlaAlaLeuGluLys 507
Db    2129 GTTCAGGTGGCTCTGGATAAG 2149

RESULT 12
US-08-612-521-1
; Sequence 1, Application US/08612521
; Patent No. 5786463
;  GENERAL INFORMATION:
;    APPLICANT:  Peery, Robert B
;    APPLICANT:  Skatrud, Paul L
;    APPLICANT:  Thornewell, Susan J
;    TITLE OF INVENTION:  MULTIPLE DRUG RESISTANCE GENE OF
;    TITLE OF INVENTION:  CRYPTOCOCCUS NEOFORMANS
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Patent Division/AEH
```

```
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  USA
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/612,521
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Hamilton, Amy E
;      REGISTRATION NUMBER:  33,894
;      REFERENCE/DOCKET NUMBER:  X-9693
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-3169
;      TELEFAX:  317-276-1294
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4224 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..4224
US-08-612-521-1

Alignment Scores:
Pred. No.:              7.87e-107     Length:         4224
Score:                  975.50        Matches:        209
Percent Similarity:     59.81%        Conservative:   114
Best Local Similarity:  38.70%        Mismatches:     180
Query Match:            37.64%        Indels:         37
DB:                     1             Gaps:           8

US-09-873-409-8 (1-514) x US-08-612-521-1 (1-4224)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    466 ATGGTCTTGGGCCTTGTGCTTGCCGTTGCAGCCGGATCTTGCCAACCTTTAATGACTTTA 525

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db    526 ATATTCGGTCGATTAACCACGTCTTTCACGAATTATGCTGTCATT---GCGAACCAAATA 582

Qy     41 SerPhePheArgLeuThr------------------------------------------ 46
Db    583 TCCCAAGGGGACTTACTCCCGAGACGTCTGCGGCTTTACAAGCAGCCAAAGATGATCTC 642

Qy     47 -----------------------LeuTyrTyrValGlyIleGlyValAlaAlaLeuIle 58
Db    643 AAGACTCAATCTGGCCACAATGCCCTTTATCTGATGGCCATTGGCATTGGAATGTTTCTA 702

Qy     59 PheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLysArgIle 78
Db    703 GCGACTTGGCTTTACATGTTCATCTGGAATGTTACTGGCGAGCTCAACTCTAAAAGGATT 762

Qy     79 ArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSerCysAsp 98
Db    763 AGAGAGCGTTACTTGGCTGCAGTGCTTAGACAGGAGATTGCCTACTTTGATGATCTGGGC 822

Qy     99 IleGlyGluLeuAsnThrArgMet---ThrAspIleAspLysIleSerAspGlyIleGly 117
Db    823 GCCGGAGAAGTCGCCACTCGCATTCAGACCGATTGTCACCTTGTCCAAGAGGGAACATCT 882

Qy    118 AspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAlaValGly 137
Db    883 GAAAAGGTCGCTCTCGTATTCCAATATGCTGGTACTTTTGTCTGCGGTTTTGTTCTAGCT 942

Qy    138 LeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIleMetAla 157
Db    943 TTTGTCCGGTCACCTCGTCTTGCGGGCGCTCTTGTTTCCATCCTGCCAGTCATCATGCTC 1002

Qy    158 SerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSerAlaTyr 177
Db   1003 TGCGGCGGTATTATGATGACGGCCATGGCCAAATATGGGACTGCGGCTCTTGATCACATC 1062

Qy    178 SerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIleAlaPhe 197
Db   1063 GCGAAAGCGGGCAGCTTGGCAGAAGAGGTCATAGGAAGTATCAGGACTGTTCAGGCTTTT 1122

Qy    198 ArgAlaGlnGluLysGluLeu---GlnArgTyrThrGlnAsnLeuLysAspAlaLysAsp 216
Db   1123 ---GGCAAGGAAAAGATCTTGGGTGACAAATTTGCCGATCACATTGAGCAGAGCAAGATT 1179

Qy    217 PheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePheMet 236
Db   1180 GTCGGTAGAAAAGGCTCCATCTTTGAAGGCTTTGGTTTGAGCATCATGTTCTTCGTCATC 1239

Qy    237 AsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGluPro 256
Db   1240 TACGCCGCTTATGCCCTCGCCTTCTTCTACGGTGGCATTCTCGTCAGTAACGGCCAAGCC 1299

Qy    257 GlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyrCys 276
Db   1300 ------GACTCCGGTATCGTCATCAACGTTTTCATGTCCATCCTCATTGGTTCCTTCTCC 1353

Qy    277 IleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPheHis 296
Db   1354 ATGGCGATGCTTGCTCCCGAGTTGGCTGCTGTCACCAAAGCGCGAGGCGCTGCTGCCAAG 1413

Qy    297 IlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyrLys 316
Db   1414 CTTTTCGCAACTATCGACCGTGTACCCGCCATTGATTCCGCCAGCGAGGAAGGTTTCAAG 1473

Qy    317 ProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSerArg 336
Db   1474 CCCGATGGCCTTCGCGGTGAGATTAGCTTTGAAAATGTCAAGTTCCATTATCCGTCCCGA 1533

Qy    337 ProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrValAla 356
Db   1534 CCCAGCATTCCTATCTTGAAGGGCTTCACTACTACGTTTGAAGCCGGTAAGACTTTCGCT 1593

Qy    357 LeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeuTyr 376
Db   1594 CTCGTCGGAGCTAGTGGAAGCGGAAAGAGTACTGTCGTTTCTCTTATTGAGAGGTTTTAC 1653

Qy    377 AspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnValArg 396
Db   1654 GACCCCGTTAGTGGTGTTGTGAAGCTTGACGGTAGAGATATCAGGTCGCTTAACCTCAAT 1713

Qy    397 HisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThrIle 416
Db   1714 TGGCTTCGTCAACAGATTGGTCTTGTATCTCAGGAACCCACTCTTTTCGGTACCACCGTC 1773

Qy    417 SerAsnAsnIleLysTyrGly---------------ArgAspAspValThrAspGluGlu 431
Db   1774 CGCGGTAACGTCGAACACGGTCTCATCGGCTCCAGATATGAGAATGCCTCTCTCGAGGAA 1833

Qy    432 ------------MetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMetGlu 447
Db   1834 AAATTTGAGCTTGTGAAGAAGGCTTGTGTCGATGCCAACGCCCATAATTTTATCATGAAA 1893

Qy    448 PheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGlyGln 467
Db   1894 CTTCCGCAAGGTTATGATACCATGGTTGGTGAGCGCGGTATGCTGTTATCTGGTGGTCAG 1953

Qy    468 LysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeuAsp 487
Db   1954 AAGCAGCGAGTTGCGATTGCTCGTGCTATCGTTTCCGACCCTAGGATTTTGTTGTTGGAC 2013
```

```
Qy    488 GluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGluLys 507
Db   2014 GAAGCTACCTCTGCTCTTGACACTCAGAGCGAAGGTATTGTGCAAGATGCTTTGGACAAG 2073

RESULT 13
US-08-612-734B-1
; Sequence 1, Application US/08612734B
; Patent No. 5914246
;  GENERAL INFORMATION:
;    APPLICANT:  Peery, Robert B.
;    APPLICANT:  Skatrud, Paul L.
;    APPLICANT:  Tobin, Matthew B.
;    TITLE OF INVENTION:  Multiple Drug Resistance Gene of
;    TITLE OF INVENTION:  Aspergillus Fumigatus
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center, DC1501
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  U.S.A.
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/612,734B
;      FILING DATE:  08-MAR-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Craig, Anne I.
;      REGISTRATION NUMBER:  32,976
;      REFERENCE/DOCKET NUMBER:  X-9681
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  617-354-9570
;      TELEFAX:  617-354-4043
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4047 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..4047
US-08-612-734B-1

Alignment Scores:
Pred. No.:              7.16e-105     Length:         4047
Score:                  959.00        Matches:        203
Percent Similarity:     59.96%        Conservative:   110
Best Local Similarity:  38.89%        Mismatches:     191
Query Match:            37.00%        Indels:         18
DB:                     2             Gaps:           6

US-09-873-409-8 (1-514) x US-08-612-734B-1 (1-4047)

Qy      1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
Db    331 ATACTCGTGAGCGCAATCTGTGCTATTGCTGCTGGTGCCGCTCTTCCCTTGTTCACAATT 390

Qy     21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
Db    391 CTCTTCGGTTCACTAGCC---TCAGCGTTCCAGGGCATATCTCTGGGTACTATGCCCTAT 447

Qy     41 SerPhePhe------ArgLeuThr---------LeuTyrTyrValGlyIleGlyValAla 55
Db    448 CATGAGTTCTATCACAAACTGACTAAGAATGTGCTTTACTTTGTGTATCTCGGTATTGCC 507

Qy     56 AlaLeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThr 75
Db    508 GAGTTTGTCACAGTCTATGTCAGCACCGTGGGTTTCATTTATACTGGCGAACATCTCACA 567

Qy     76 LysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAsp 95
Db    568 CAGAAGATCCGTGAAAATTATCTTGAGGCTATCCTGAGGCAGAATATGGCTTACTTCGAC 627

Qy     96 SerCysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAsp 114
Db    628 AAGTTGGGCGCCGGTGAAGTTACCACGCGTATCACTGCTGATACCAACCTGATCCAGGAC 687

Qy    115 GlyIleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeu 134
Db    688 GCCATCTCTGAGAAAGTTGGTCTCACTTTGACCGCATTCGCCACATTTGTAACCGCATTT 747

Qy    135 AlaValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeu 154
Db    748 ATTGTCGCCTACGTCAAGTATTGGAAGTTGGCTCTGATCTGTACCTCAACTATCGTCGCG 807

Qy    155 IleMetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeu 174
Db    808 CTGGTCATGGTTATGGGAGGTGGGTCGAGGTTTATTGTGAAGTACAGCAAGAAATCTATT 867

Qy    175 SerAlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrVal 194
Db    868 GAAAGTTATGGTGCTGGTGGAACTGTCGCGGAAGAAGTCATCAGCTCCATTCGGAATGCT 927

Qy    195 IleAlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAla 214
Db    928 ACCGCTTTCGGCACTCAGGATAAGCTCGCCAAGCAATACGAAACCCATCTGGCTGAGGCT 987

Qy    215 LysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhe 234
Db    988 GAAAAATGGGGCGTCAAACAACAGGTCATCCTTGGTATGATGATTGGTGGTATGTTCGGT 1047

Qy    235 PheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGly 254
Db   1048 ATCATGTTCTCGAACTATGGTCTCGGTTTCTGGATGGGATCTCGCTTCGTTGTCGGTAAA 1107

Qy    255 GluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSer 274
Db   1108 GAA------GTCAACGTGGGCCAAGTTCTGACAGTTTTGATGTCTATCCTGATCGGTTCG 1161

Qy    275 TyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAla 294
Db   1162 TTCAGTTTGGGCAACGTCGCCCCCAATGGTCAGGCCTTTACGAATGGTGTTGCTGCGGCC 1221

Qy    295 PheHisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGly 314
Db   1222 GCGAAGATTTACAGCACGATTGACCGCAGATCGCCACTGGACCCCTATTCTGACGAAGGG 1281

Qy    315 TyrLysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrPro 334
Db   1282 AAGGTACTCGACCATTTTGAAGGAAATATCGAATTTCGCAATGTCAAACACATCTACCCT 1341

Qy    335 SerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThr 354
Db   1342 TCAAGACCCGAAGTTACAGTCATGGAAGATGTCTCTTTATCGATGCCCGCCGGAAAGACT 1401

Qy    355 ValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArg 374
Db   1402 ACCGCATTGGTGGGCCCATCTGGCTCTGGAAAGAGTACTGTTGTCGGCTTGGTGGAGCGC 1461

Qy    375 LeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsn 394
Db   1462 TTTTACCTTCCAGTAGGAGGCCAGGTATTGCTGGACGGCCATGATATCCAAACCCTCAAC 1521

Qy    395 ValArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThr 414
Db   1522 CTCCGTTGGCTGCGACAGCAGATCTCTCTTGTCAGCCAGGAACCTGTTCTTTTCAGCACC 1581

Qy    415 ThrIleSerAsnAsnIleLysTyrGly--------------------------ArgAsp 425
```

```
Db     1582 ACGATCTTTAGAAACATCGAACATGGCTTGATTGGCACCAAATTCGAGCATGAGTCGAAG 1641

Qy      426 AspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIle 445
            |||         |||   :::|||   ||||||||   |||||||||:::||||||||
Db     1642 GACAAGATTAGAGAGCTCGTTGAGAATGCGGCCAGAATGGCCAATGCTCATGATTTTATT 1701

Qy      446 MetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGly 465
            |||      |||      ::::::|||   ||||||::::::|||      :::||||||
Db     1702 ATGGCTCTGCCTGAAGGTTACGATACGAATGTGGGTCAGCGTGGTTTCTTACTTTCAGGA 1761

Qy      466 GlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIle 485
            |||||||||||||||||||||||||||||||||:::|||   :::||||||||||||:::
Db     1762 GGTCAGAAGCAACGTATTGCCATTGCTCGTGCCATTGTCAGTGACCCCAAGATTCTGTTG 1821

Qy      486 LeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeu 505
            ||||||||||||||||||||||||||||||:::::|||:::      ||||||||||||||
Db     1822 CTTGATGAAGCTACATCAGCTTTGGATACCAAGTCCGAGGGCGTCGTCCAAGCCGCTCTT 1881

Qy      506 GluLys 507
            :::|||
Db     1882 GATAAA 1887
```

RESULT 14
US-08-612-734B-3
; Sequence 3, Application US/08612734B
; Patent No. 5914246
; GENERAL INFORMATION:
;   APPLICANT: Peery, Robert B.
;   APPLICANT: Skatrud, Paul L.
;   APPLICANT: Tobin, Matthew B.
;   TITLE OF INVENTION: Multiple Drug Resistance Gene of
;   TITLE OF INVENTION: Aspergillus Fumigatus
;   NUMBER OF SEQUENCES: 9
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Eli Lilly and Company
;   STREET: Lilly Corporate Center, DC1501
;   CITY: Indianapolis
;   STATE: Indiana
;   COUNTRY: U.S.A.
;   ZIP: 46285
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/612,734B
;   FILING DATE: 08-MAR-1996
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Craig, Anne I.
;   REGISTRATION NUMBER: 32,976
;   REFERENCE/DOCKET NUMBER: X-9681
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 617-354-9570
;   TELEFAX: 617-354-4043
; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 4800 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;   MOLECULE TYPE: other nucleic acid
US-08-612-734B-3

Alignment Scores:

Pred. No.:	2.16e-102	Length:	4800
Score:	939.50	Matches:	202
Percent Similarity:	58.35%	Conservative:	102
Best Local Similarity:	38.77%	Mismatches:	182
Query Match:	36.25%	Indels:	35
DB:	2	Gaps:	6

US-09-873-409-8 (1-514) x US-08-612-734B-3 (1-4800)

```
Qy        1 MetIleLeuGlyIleLeuAlaSerLeuValAsnGlyAlaCysLeuProLeuMetProLeu 20
            ::::::   |||   |||||||||            |||      |||      ||||||
Db      693 ATTCTCTTCGGTTCACTAGCCTCAGCGTTCCAGGGCATATCTCTGGGTACTATGCCCTAT 752

Qy       21 ValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeuValGlnThrAsnThrTyr 40
                                                                     :::
Db      753 ---------------------------------------------------------CAT 755

Qy       41 SerPhePhe---ArgLeuThr---------LeuTyrTyrValGlyIleGlyValAlaAla 56
               |||:::   :::||||||         ||||||:::|||   :::|||:::|||
Db      756 GAGTTCTATCACAAACTGACTAAGAATGTGCTTTACTTGTGTATCTCGGTATTGCCGAG 815

Qy       57 LeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThrAlaAlaArgGlnThrLys 76
               :::      |||:::                :::|||   |||            |||:::
Db      816 TTTGTCACAGTCTATGTCAGCACCGTGGGTTTCATTTATACTGGCGAACATCTCACACAG 875

Qy       77 ArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAspIleGlyTrpPheAspSer 96
            :::||||||:::   :::      ::::::|||   |||::::::   :::||||||
Db      876 AAGATCCGTGAAAATTATCTTGAGGCTATCCTGAGGCAGAATATGGCTTACTTCGACAAG 935

Qy       97 CysAspIleGlyGluLeuAsnThrArgMetThr---AspIleAspLysIleSerAspGly 115
                        ||||||:::   ||||||:::|||   |||   :::   |||   |||
Db      936 TTGGGCGCCGGTGAAGTTACCACGCGTATCACTGCTGATACCAACCTGATCCAGGACGCC 995

Qy      116 IleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThrPheSerIleGlyLeuAla 135
            |||   :::|||:::   |||               :::||||||
Db      996 ATCTCTGAGAAAGTTGGTCTCACTTTGACCGCATTCGCCACATTTGTAACCGCATTTATT 1055

Qy      136 ValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeuSerThrSerProLeuIle 155
            |||         ||||||   |||||||||   |||:::         ||||||            :::
Db     1056 GTCGCCTACGTCAAGTATTGGAAGTTGGCTCTGATCTGTACCTCAACTATCGTCGCGCTG 1115

Qy      156 MetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeuThrSerLysGluLeuSer 175
            :::                     ||||||   ::::::      :::   |||   :::
Db     1116 GTCATGGTTATGGGAGGTGGGTCGAGGTTTATTGTGAAGTACAGCAAGAAATCTATTGAA 1175

Qy      176 AlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSerSerIleArgThrValIle 195
            :::|||         |||   |||||||||||||||:::||||||||||||
Db     1176 AGTTATGGTGCTGGTGGAACTGTCGCGGAAGAAGTCATCAGCTCCATTCGGAATGCTACC 1235

Qy      196 AlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGlnAsnLeuLysAspAlaLys 215
            ||||||      |||:::|||      ::::::|||      :::|||   :::|||:::
Db     1236 GCTTTCGGCACTCAGGATAAGCTCGCCAAGCAATACGAAACCCATCTGGCTGAGGCTGAA 1295

Qy      216 AspPheGlyIleLysArgThrIleAlaSerLysValSerLeuGlyAlaValTyrPhePhe 235
            :::|||:::|||:::   :::      :::      :::|||      ::::::
Db     1296 AAATGGGGCGTCAAACAACAGGTCATCCTTGGTATGATGATTGGTGGTATGTTCGGTATC 1355

Qy      236 MetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSerLeuIleLeuAsnGlyGlu 255
            |||         |||||||||   ||||||   |||:::         ::::::      |||
Db     1356 ATGTTCTCGAACTATGGTCTCGGTTTCTGGATGGGATCTCGCTTCGTTGTCGGTAAAGAA 1415

Qy      256 ProGlyTyrThrIleGlyThrValLeuAlaValPhePheSerValIleHisSerSerTyr 275
                     :::|||   ||||||   |||      |||::::::      |||:::
Db     1416 ------GTCAACGTGGGCCAAGTTCTGACAGTTTTGATGTCTATCCTGATCGGTTCGTTC 1469

Qy      276 CysIleGlyAlaAlaValProHisPheGluThrPheAlaIleAlaArgGlyAlaAlaPhe 295
               :::|||            |||:::   :::   |||            ||||||
Db     1470 AGTTTGGGCAACGTCGCCCCCAATGGTCAGGCCTTTACGAATGGTGTTGCTGCGGCCGCG 1529

Qy      296 HisIlePheGlnValIleAspLysLysProSerIleAspAsnPheSerThrAlaGlyTyr 315
               |||:::      ||||||::::::      :::|||   :::|||      |||
Db     1530 AAGATTTACAGCACGATTGACCGCAGATCGCCACTGGACCCCTATTCTGACGAAGGGAAG 1589

Qy      316 LysProGluSerIleGluGlyThrValGluPheLysAsnValSerPheAsnTyrProSer 335
               :::      ||||||   :::||||||:::||||||         |||||||||
Db     1590 GTACTCGACCATTTTGAAGGAAATATCGAATTTCGCAATGTCAAACACATCTACCCTTCA 1649

Qy      336 ArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIleLysSerGlyGluThrVal 355
            ||||||   :::   :::::::::   ::::::|||   :::   :::|||:::|||
Db     1650 AGACCCGAAGTTACAGTCATGGAAGATGTCTCTTTATCGATGCCCGCCGGAAAGACTACC 1709
```

```
Qy     356 AlaLeuValGlyLeuAsnGlySerGlyLysSerThrValValGlnLeuLeuGlnArgLeu 375
Db    1710 GCATTGGTGGGCCCATCTGGCTCTGGAAAGAGTACTGTTGTCGGCTTGGTGGAGCGCTTT 1769

Qy     376 TyrAspProAspAspGlyPheIleMetValAspGluAsnAspIleArgAlaLeuAsnVal 395
Db    1770 TACCTTCCAGTAGGAGGCCAGGTATTGCTGGACGGCCATGATATCCAAACCCTCAACCTC 1829

Qy     396 ArgHisTyrArgAspHisIleGlyValValSerGlnGluProValLeuPheGlyThrThr 415
Db    1830 CGTTGGCTGCGACAGCAGATCTCTCTTGTCAGCCAGGAACCTGTTCTTTTCAGCACCACG 1889

Qy     416 IleSerAsnAsnIleLysTyrGly--------------------------ArgAspAsp 426
Db    1890 ATCTTTAGAAACATCGAACATGGCTTGATTGGCACCAAATTCGAGCATGAGTCGAAGGAC 1949

Qy     427 ValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsnAlaTyrAspPheIleMet 446
Db    1950 AAGATTAGAGAGCTCGTTGAGAATGCGGCCAGAATGGCCAATGCTCATGATTTTATTATG 2009

Qy     447 GluPheProAsnLysPheAsnThrLeuValGlyGluLysGlyAlaGlnMetSerGlyGly 466
Db    2010 GCTCTGCCTGAAGGTTACGATACGAATGTGGGTCAGCGTGGTTTCTTACTTTCAGGAGGT 2069

Qy     467 GlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsnProLysIleLeuIleLeu 486
Db    2070 CAGAAGCAACGTATTGCCATTGCTCGTGCCATTGTCAGTGACCCCAAGATTCTGTTGCTT 2129

Qy     487 AspGluAlaThrSerAlaLeuAspSerGluSerLysSerAlaValGlnAlaAlaLeuGlu 506
Db    2130 GATGAAGCTACATCAGCTTTGGATACCAAGTCCGAGGGCGTCGTCCAAGCCGCTCTTGAT 2189

Qy     507 Lys 507
Db    2190 AAA 2192
```

```
RESULT 15
US-08-996-545-1
; Sequence 1, Application US/08996545
; Patent No. 5928898
;  GENERAL INFORMATION:
;    APPLICANT:  Skatrud, Paul L.
;    APPLICANT:  de Waard, Maarten A.
;    APPLICANT:  Peery, Robert B.
;    APPLICANT:  Andrade, Alan C.
;    TITLE OF INVENTION:  Multiple Drug Resistance Gene atrD of
;    TITLE OF INVENTION:  Aspergillus Nidulans
;    NUMBER OF SEQUENCES:  3
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Eli Lilly and Company
;      STREET:  Lilly Corporate Center
;      CITY:  Indianapolis
;      STATE:  Indiana
;      COUNTRY:  U.S.
;      ZIP:  46285
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/996,545
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Webster, Thomas D.
;      REGISTRATION NUMBER:  39,872
;      REFERENCE/DOCKET NUMBER:  X-11766
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  317-276-3334
;      TELEFAX:  317-276-2763
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  4002 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  1..4002
US-08-996-545-1

Alignment Scores:
Pred. No.:                2.75e-102      Length:         4002
Score:                    937.50         Matches:        201
Percent Similarity:       59.96%         Conservative:   103
Best Local Similarity:    39.64%         Mismatches:     181
Query Match:              36.17%         Indels:         23
DB:                       2              Gaps:           5

US-09-873-409-8 (1-514) x US-08-996-545-1 (1-4002)
```

```
Qy      15 LeuProLeuMetProLeuValLeuGlyGluMetSerAspAsnLeuIleSerGlyCysLeu 34
Db     368 GTGCCATTGCTGCCGCGTCGACTTTCCAGA-----------------GGATAATGTTAT 409

Qy      35 ValGlnThrAsnThrTyrSerPhePhe-----------ArgLeuThrLeuTyrTyrVal 50
Db     410 ATCAAATCTCGTACGAC-GAGTTCTATGATGAATTGACCAAGAACGTACTGTACTTCGTA 468

Qy      51 GlyIleGlyValAlaAlaLeuIlePheGlyTyrIleGlnIleSerLeuTrpIleIleThr 70
Db     469 TACCTCGGTATCGGCGAGTTTGTCACTGTCTATGTTAGTACTGTTGGCTTCATCTATACC 528

Qy      71 AlaAlaArgGlnThrLysArgIleArgLysGlnPhePheHisSerValLeuAlaGlnAsp 90
Db     529 GGAGAACACGCCACGCAGAAGATCCGCGAGTATTACCTTGAGTCTATCCTGCGCCAGAAC 588

Qy      91 IleGlyTrpPheAspSerCysAspIleGlyGluLeuAsnThrArgMetThr---AspIle 109
Db     589 ATTGGCTATTTTGATAAACTCGGTGCCGGGGAAGTGACCACCCGTATAACAGCCGATACA 648

Qy     110 AspLysIleSerAspGlyIleGlyAspLysIleAlaLeuLeuPheGlnAsnMetSerThr 129
Db     649 AACCTTATCCAGGATGGCATTTCGGAGAAGGTCGGTCTCACTTTGACTGCCCTGGCGACA 708

Qy     130 PheSerIleGlyLeuAlaValGlyLeuValLysGlyTrpLysLeuThrLeuValThrLeu 149
Db     709 TTCGTGACAGCATTCATTATCGCCTACGTCAAATACTGGAAGTTGGCTCTAATTTGCAGC 768

Qy     150 SerThrSerProLeuIleMetAlaSerAlaAlaAlaCysSerArgMetValIleSerLeu 169
Db     769 TCAACAATTGTGGCCCTCGTTCTCACCATGGGCGGTGGTTCTCAGTTTATCATCAAGTAC 828

Qy     170 ThrSerLysGluLeuSerAlaTyrSerLysAlaGlyAlaValAlaGluGluValLeuSer 189
Db     829 AGCAAAAAGTCGCTTGACAGCTACGGTGCAGGCGGCACTGTTGCGGAAGAGGTCATCAGC 888

Qy     190 SerIleArgThrValIleAlaPheArgAlaGlnGluLysGluLeuGlnArgTyrThrGln 209
Db     889 TCCATCAGAAATGCCACAGCGTTTGGCACCCAAGACAAGCTTGCGAAGCAGTATGAGGTC 948

Qy     210 AsnLeuLysAspAlaLysAspPheGlyIleLysArgThrIleAlaSerLysValSerLeu 229
Db     949 CACTTAGACGAAGCTGAGAAATGGGGAACAAAGAACCAGATTGTCATGGGTTTCATGATT 1008

Qy     230 GlyAlaValTyrPhePheMetAsnGlyThrTyrGlyLeuAlaPheTrpTyrGlyThrSer 249
Db    1009 GGCGCCATGTTTGGCCTTATGTACTCGAACTACGGTCTTGGCTTCTGGATGGGTTCTCGT 1068

Qy     250 LeuIleLeuAsnGlyGluProGlyTyrThrIleGlyThrValLeuAlaValPhePheSer 269
```

```
Db   1069 TTCCTGGTAGATGGT------GCAGTCGATGTGGGTGATATTCTCACAGTTCTCATGGCC 1122

Qy    270 ValIleHisSerSerTyrCysIleGlyAlaAlaValProHisPheGluThrPheAlaIle 289
          ::::::       |||:::   :::|||          |||:::   :::   |||
Db   1123 ATCTTGATCGGATCGTTCTCCTTGGGGAACGTTAGTCCAAATGCTCAAGCATTTACAAAC 1182

Qy    290 AlaArgGlyAlaAlaPheHisIlePheGlnValIleAspLysLysProSerIleAspAsn 309
          |||      ||||||      ||||||     ||||||::::::       :::|||
Db   1183 GCTGTGGCCGCGGCCGCAAAGATATTTGGAACGATCGATCGCCAGTCCCCATTAGATCCA 1242

Qy    310 PheSerThrAlaGlyTyrLysProGluSerIleGluGlyThrValGluPheLysAsnVal 329
          :::|||       |||      :::       ||||||   :::|||   :::||||||
Db   1243 TATTCGAACGAAGGGAAGACGCTCGACCATTTTGAGGGCCACATTGAGTTACGCAATGTC 1302

Qy    330 SerPheAsnTyrProSerArgProSerIleLysIleLeuLysGlyLeuAsnLeuArgIle 349
                   |||||||||||||||   :::   :::::::::   ::::::|||   :::
Db   1303 AAGCATATTTACCCATCTAGACCCGAGGTCACCGTCATGGAGGATGTTTCTCTGTCAATG 1362

Qy    350 LysSerGlyGluThrValAlaLeuValGlyLeuAsnGlySerGlyLysSerThrValVal 369
             :::|||:::|||   ||||||||||||   :::||||||||||||||||||||||||
Db   1363 CCCGCTGGAAAAACAACCGCTTTAGTCGGCCCCTCTGGCTCTGGAAAAAGTACGGTGGTC 1422

Qy    370 GlnLeuLeuGlnArgLeuTyrAspProAspAspGlyPheIleMetValAspGluAsnAsp 389
             |||::::::|||   |||   |||      |||   :::::::::|||   :::|||
Db   1423 GGCTTGGTTGAGCGATTCTACATGCCTGTTCGCGGTACGGTTTTGCTGGATGGCCATGAC 1482

Qy    390 IleArgAlaLeuAsnValArgHisTyrArgAspHisIleGlyValValSerGlnGluPro 409
          |||:::   ||||||:::|||         |||      |||   :::|||||||||||||||
Db   1483 ATCAAGGACCTCAATCTCCGCTGGCTTCGCCAACAGATCTCTTTGGTTAGCCAGGAGCCT 1542

Qy    410 ValLeuPheGlyThrThrIleSerAsnAsnIleLysTyrGly------------------ 423
          |||||||||||||||||||||         ||||||::::::|||
Db   1543 GTTCTTTTTGGCACGACGATTTATAAGAATATTAGGCACGGTCTCATCGGCACAAAGTAC 1602

Qy    424 ---------ArgAspAspValThrAspGluGluMetGluArgAlaAlaArgGluAlaAsn 440
                   :::|||         |||   :::|||   ||||||:::   ||||||
Db   1603 GAGAATGAATCCGAGGATAAGGTCCGGGAACTCATCGAGAACGCGGCAAAAATGGCGAAT 1662

Qy    441 AlaTyrAspPheIleMetGluPheProAsnLysPheAsnThrLeuValGlyGluLysGly 460
          |||:::|||||||||         |||      :::   |||   ||||||::::::|||
Db   1663 GCTCATGACTTTATTACTGCCTTGCCTGAAGGTTATGAGACCAATGTTGGGCAGCGTGGC 1722

Qy    461 AlaGlnMetSerGlyGlyGlnLysGlnArgIleAlaIleAlaArgAlaLeuValArgAsn 480
                :::|||||||||||||||||||||||||||||||||||||||:::|||   :::
Db   1723 TTTCTCCTTTCAGGTGGCCAGAAACAGCGCATTGCAATCGCCCGTGCCGTTGTTAGTGAC 1782

Qy    481 ProLysIleLeuIleLeuAspGluAlaThrSerAlaLeuAspSerGluSerLysSerAla 500
          ||||||||||||:::||||||||||||||||||||||||||||::::::|||:::
Db   1783 CCAAAAATCCTGCTCCTGGATGAAGCTACTTCGGCCTTGGACACAAAATCCGAAGGCGTG 1842

Qy    501 ValGlnAlaAlaLeuGluLys 507
          ||||||||||||||||||:::
Db   1843 GTTCAAGCAGCTTTGGAGAGG 1863
```

Search completed: March 31, 2003, 15:38:51
Job time : 138.273 secs